3-21-2013

Control Variates and Optimal Designs in Metamodeling

Joshua B. Meents

Control Variates and Optimal Designs in Metamodeling

THESIS

Joshua B. Meents, Captain, USAF

AFIT-ENS-13-M-11

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT-ENS-13-M-11

Control Variates and Optimal Designs in Metamodeling

THESIS

Presented to the Faculty of the School of Engineering

of the Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Operations Research

Joshua B. Meents, BS

Captain, USAF

March 2012

AFIT-ENS-13-M-11

Control Variates and Optimal Designs in Metamodeling

Joshua B. Meents, BS
Captain, USAF

Approved:

Mark A. Friend, Lt Col, USAF (Chairman)	Date
Kenneth Bauer, PhD, USAF (Member)	Date

AFIT-ENS-13-M-11

Abstract

At the heart of most modeling issues is a focus on variance reduction. Experimental designs are chosen based on both efficiency and a variety of variance based criteria. In many situations due to cost, time and availability issues it is beneficial to produce metamodels of simulations. Experimental designs for the region of operability are constructed to collect the simulation output required to construct representative metamodels. Independently, the method of control variates is a well established technique often employed to reduce variance in discrete event simulations. This thesis explores the variance reduction benefits that can be obtained by combining optimal experimental designs with control variates in multipopulation simulation experiments when constructing simulation metamodels. A variety of variance measures of effectiveness are used to demonstrate the theoretical benefits obtained by this approach. In addition, a randomly selected data set from within the design region is used to demonstrate the practical application and reduction of predictive variance obtained using this methodology.

Acknowledgments

I would like to express my sincere appreciation to my faculty advisor, Lt. Col. Mark Friend, for his guidance and support throughout the course of this thesis effort. The insight and experience was certainly appreciated.

Joshua B. Meents

Table of Contents

List of Figures

List of Tables

List of Equations

Control Variates and Optimal Designs in Metamodeling

1. Introduction

1.1 Problem Background

As computing power increases rapidly, the number of potential systems that can be modeled with accuracy increases as well. However, every real world situation has variance throughout the process and this variance must be captured by simulations of real world processes with distribution functions. This leads the analysis team to results that involve a range of values in order to account for the aggregation of variance and a multitude of potential scenarios. Some models, even with current computing power take considerable time and resources to design and run and still more resources are required to analyze the model's output. By using these initial simulations to characterize the model and create a metamodel, we can provide estimates without rerunning the model. By creating better estimators and reduced variance within the output, the simulation creates better results for the analysis team to then follow.

1.2 Purpose of the Study

The purpose of this thesis is to show the potential benefits of using control variates when using optimal designs to create metamodels from simulation models. These benefits could include increasing the prediction accuracy and reducing the variance of the output with fewer runs and less resources. In experimental design there are a variety of variance optimal design techniques based on different design criteria. This study will seek if additional variance reduction might be garnered through the application of control variates.

1.3 Questions to be Investigated

Several questions will be investigated by this research into the application of control variates across different optimal designs and within a real world model. The most important question to be answered is whether control variates do indeed provide a benefit to an analyst seeking to create a metamodel from a simulation and how significant is the benefit. The next question is whether or not this benefit is observed only with specific design structures or may be applied in all cases. Finally, the study will determine if control variate models create better predictions for points in the design space than non-control variate models.

1.4 Hypotheses

My hypothesis to these questions is that control variates, when used properly, will have a clear benefit in prediction as well as decreasing the half-widths and variance estimates. I hypothesize that this benefit will be seen through all optimal design structures when compared to their non-control variate case. I also hypothesize that when different designs are compared, the control variate and non-control variate models will perform similarly. For example, if the non control variate model is better than another design's non control variate model, then the same should happen when control variates are applied. Within these hypotheses I predict that changes to the model will impact the benefit of control variates. Therefore, the model must be correct to begin with and control variates implemented correctly before investigating the interaction between control variates and different optimal designs.

1.5 Rationale and Theoretical Framework

The rationale behind my hypotheses is that control variates and optimal designs have both shown to assist in metamodeling analysis as seen in the forthcoming literature review. The techniques do not contradict each other and the use of one does not negatively impact the benefits gained by using the other. Therefore, by using them together the benefits of each should still be visible. Theoretically, control variates allow the analyst to use the true values that occur within the simulation to assist in interpreting the result while an optimal design allows the analyst to investigate the design space as efficiently as possible. Therefore, using both methods allows the analyst to get the best possible picture of the design space and then adjust these results with control variates to create the best model possible.

1.6 Assumptions

Several assumptions exist within the use of control variates and optimal designs as well as various output analysis techniques that will be used. These will be discussed in detail later. For now, overall assumptions for the thesis include the use of a verified and validated model, assumptions of the individual techniques used are true, and that the reader has a basic knowledge of simulation modeling and statistical analysis.

1.7 Importance of Study

Insights gained by this thesis research will benefit analyses that seek to create metamodels to accurately characterize simulations. The reduction in unexplained variance through the use of control variates will allow for increased precision in these metamodels without any increase in the number of experimental design points. In a world where everything needs to be faster and cheaper, this can be very advantageous to anyone conducting computer simulations that meet the very limited needs of optimal designs and control variates.

1.8 Definition of Terms

As mentioned in the assumptions, this thesis is written assuming the reader has a basic understanding of metamodeling and statistical analysis. Generally speaking, metamodeling is a model of a model where the construction and development of the rules, constraints, models and theories are useful for modeling a predefined class of problems. Control variates is a technique which uses observed randomly distributed variable values in a model to create a better estimate of the response by relying on the correlation between the control variate and the response. The factors which are used as control variates will be labeled control variates, and will be represented by a q in the equations. The factors of interest, which are changed throughout a design to see their impact on the response, will be represented with the variable p. The number of responses of interest from the model will be represented with the variable m. The number of runs within a single replication is represented with the variable k, while the number of replications for a design is l. Optimal design construction uses a measure of efficiency to create a design within model constraints that is optimal with respect to the chosen measure. There are several measures of efficiency and therefore for any given number of factors and observations several different optimal designs, each with their own advantages and disadvantages may be constructed. Several of these experimental design construction methods will be described in the literature review.

Table 1: The variables used in future equations.

Term	Meaning
q	Number of control variates
p	Number of factors of interest
m	Number of responses
k	Number of runs per replication
l	Number of replications

1.9 Scope and Delimitations

There are many widespread assumptions throughout statistical analysis. This research will not seek to prove or disprove them, but instead will accept the prior work and proofs as accurate. The scope of this project is to show the impact of combining these various reduction techniques for the models shown so that they may be considered for use by others. The scope of this research does not limit their benefit to only these scenarios; however it does not prove it beneficial in a universal case either. Complex simulations can have what seems like a countless number of moving parts and the benefit from these variance reduction techniques may not be easily seen as the system becomes ever more complex. Meanwhile, due to the low cost to implement these techniques and the potential benefits to a research team in the instances it does work, control variates should be considered when constructing metamodels from simulation data.

1.10 Preview

This thesis explores and discusses prior research in chapter 2, the literature review section. This chapter clearly explains the statistical concepts and methods used in the thesis such as control variates, optimal experimental designs, analysis of covariance, evaluation techniques, and metamodeling. This provides a strong basis of knowledge for the reader to then explore the methodology used in this research.

Chapter 3, the methodology section, details the process used to create and analyze two different simulations of interest. The first simulation analyzed is an expansion of a simulation developed by Arnold, Nozari, and Pegden (1984), of the waiting time experienced by cars on a single lane section of road. The second simulation is a real world adaptation from an Air Force unit that is responsible for mission planning. The statistical methods used to analyze the benefits of control variates with these models is primarily based on prior research of Arnold, Nozari, and Pegden (1984) and Porta Nova and Wilson (Nov. 1989).

Following the methodology chapter, the analysis of the effectiveness of control variates using these two models is discussed in detail in chapter 4. Potential benefits of control variates when creating metamodels using different experimental designs are demonstrated to assist the potential research team employ control variates effectively. In order to measure these benefits, each model created is evaluated using several variance estimators as well as the mean squared error for predicting several randomly selected points and the corresponding prediction half width.

Chapter 5 concludes this thesis with a synopsis of the research, the significance of the research and results found, and recommendations for further research.

2. Literature Review

2.1 Chapter Overview

This section describes prior research in control variates, metamodeling, analysis of covariance, optimal designs, and some statistical evaluative methods. The goal of this discussion is to demonstrate what research has been conducted in the past while giving a basic understanding of how each method is used in current research. This section also provides references to published journals and textbooks where the interested reader can find additional information on the topics discussed. The methodology found in chapter 3 is based on this past research in many cases, making understanding of this foundational material very important. Any reader who already has a full understanding of these areas can go to chapter 3 to look at the specific methodology used in this research.

2.2 Control Variates

2.2.1 Overview

Prior research has been done in many different scenarios and applications of control variates as well as optimal experimental designs. Control variate studies include scenarios with univariate and multivariate responses, control variates, and factor settings as well as known and unknown variance cases. Single population research, where a simulation is conducted with only one setting of factor levels, has been conducted by Lavenberg et al (1978), Kleijnen (1974), and Cheng (1978). Cheng (1978) also assumes the variance is known for the control variates throughout his research. Lavenberg et al (1978) developed the use of control variates for single response simulations. Rubenstein and Marcus (1981) extended the research of single population simulations to computer experiments that include multivariate responses. From here, Arnold, Nozari, and Pegden (1984), and Porta Nova and Wilson (1989) expand from the single population to examine multiple population scenarios. Arnold, Nozari, and Pegden (1984) look into only single response scenarios when the variance is known and when the variance is unknown. Porta Nova and Wilson (1989) took this start and expanded it to scenarios with multivariate responses. Both articles investigate univariate and multivariate factors and control variates. All of these research journals cover the variance reduction benefits of control variates and their ease of use. However, while Arnold, Nozari, and Pegden (1984) and Porta Nova and Wilson (1989) mention multipopulation test designs, they never investigate the added benefit of different experimental design structures. This thesis will expand on this prior research and find whether the optimal designs, known throughout Design of Experiments to reduce variance in the prediction model, can be combined with control variates to create an even better metamodel with even less variance.

Variation is at the center of any process. It is what makes it so difficult to predict every potential scenario. Simulations use distribution functions to more accurately capture the real-life variation which exists. The resultant output, which is a function of random variables, is itself a random variable. There are many ways to account for variation in a simulation to attempt to increase the precision of results and obtain smaller confidence intervals from the simulation. These variance reduction techniques include common random numbers (CRN), antithetic variates (AV), and control variates (CV). Control variates are the method of choice for this paper to improve the results of different simulations.

2.2.2 Single Control Variates

Control variates attempt to use the correlation between the known distributions used to create the model, and the measure of interest that the simulation may return (Law, 2007). For

Table 2: Areas of prior research on control variates.

	Population		Factors		Variance		Response		Design	
Author	Single	Multiple	Single	Multiple	Known	Unknown	Single	Multiple	Single	Multiple
Arnold (1984)		X		X	X	X	X		X	
Cheng (1978)	X			X	X		X		X	
Kleijnen (1974)	X		X			X	X		X	
Lavenberg (1978)	X			X		X	X		X	
Marcus (1981)	X			X		X		X	X	
Porta Nova (1989)		X		X		X		X	X	
This Thesis		X		X		X	X	X		X

example, if a simulation is run to model a serving process, the waiting time might be the unknown variable the experimenter wants to find while the service and interarrival times are used to create the model. Because we know the service and arrival rates, as they were input into the computer based on data, we can use their known values to make adjustments to the simulations waiting time. This adjustment should move the simulation value towards the true value. In this example it can become quite intuitive, an increase in service time would be directly related to the waiting time, making us more confident on the true waiting time by knowing the service time and creating a reduction on the half width that surrounds the waiting time estimation. On the contrary, an increase in interarrival times would give servers more time to serve customer and equate to a decrease in waiting time. This suspected inverse relationship can be applied to the simulation output by adjusting the result towards a more accurate number and reduce the variance of the confidence interval.

The equation for applying a control variate to predict the output variable, Y_a, can go as follows (Law, 2007):

$$Y_a = Y - a(C - v) \qquad \text{Equation 1}$$

Y: estimate the simulation returns for the value of interest
a: multiplier for the CV (positive value when directly correlated)
C : value the simulation returns for the CV
v: is the known expectation of C

Therefore, Y decreases when C is greater than its known expectation or is adjusted up when it is less than its known expectation. The addition of these values does not change the expected value of Y as shown in Equation 2 (Law, 2007):

$$E[Y] = \mu \text{ and } E[C] = v; \text{ so} \qquad \text{Equation 2}$$
$$E[Y_a] = E[Y - a(C - v)] = E[Y] - a\{E[C] - E[v]\} = E[Y] - a(v - v) = E[Y]$$

The control variate also has the effect of reducing the variance of Y as long as certain conditions are met:

$$\mathrm{Var}(Y_a) = \mathrm{Var}(Y) + a^2\mathrm{Var}(C) - 2a\mathrm{Cov}(Y, C) \qquad \text{Equation 3}$$

As shown in Equation 3 there will be a reduction in variance as long as $a^2\mathrm{Var}(C) < 2a\mathrm{Cov}(Y, C)$ (Law, 2007). Equation 3 shows why the choice of *C* and *a* are very important to the effectiveness of variance reduction from the CV. It may seem easy to place *a* at ±1, however this puts the entire benefit on the choice of C alone (Law 2007). By adjusting *a* as well, we can improve our estimation even more. However, since we do not know the true value for the Cov(Y, C); methods must be used to estimate it and then find the best value of *a* as to maximize the variance reduction. Because the best value for *a* should be when the derivative of the variance is set to zero, then solve for *a* we find,

$$\frac{d(\mathrm{Var}(Y))}{da} = 2a\mathrm{Var}(C) - 2\mathrm{Cov}(Y, C) = 0, \text{then } a^* = \frac{\mathrm{Cov}(Y, C)}{\mathrm{Var}(C)} \text{ (Law, 2007)} \qquad \text{Equation 4}$$

By substituting this value back into the equation for $\mathrm{Var}(X_a)$ we obtain an estimate of $\mathrm{Var}(Y_a^*)$,

$$\mathrm{Var}(Y_a^*) = \mathrm{Var}(Y) - \frac{\mathrm{Cov}(Y,C)^2}{\mathrm{Var}(C)} = \left(1 - {p_{xy}}^2\right)\mathrm{Var}(Y) \quad \text{(Law, 2007)} \qquad \text{Equation 5}$$

Where ${p_{xy}}^2$ is the correlation between C and Y. From here it is easy to see that any correlation between C and Y would result in a decrease of variance on Y and a correlation of 1 would mean that C and Y are completely correlated and C could be used to predict Y perfectly every time, eliminating all of the variance (Law, 2007).

As mentioned before, because the true value of Y, and therefore the true value of $\mathrm{Cov}(Y, C)$, are unknown, we cannot just fill in these equations with our simulation values. This makes the user have to estimate the best value for a*. This can be done using the information from the simulation. Since we will have replications of the simulation, we can use these samples to estimate the sample $\mathrm{Cov}(Y, C), \hat{C}_{xy}$, and then â:

$$\hat{C}_{xy}(n) = \frac{\sum_{j=1}^{n}\left[Y_j - \bar{Y}(n)\right]\left[C_j - \bar{C}(n)\right]}{n-1} \text{ and } \hat{a}^*(n) = \frac{\hat{C}_{cy}}{S_C^2(n)} \quad \text{(Law, 2007)} \qquad \text{Equation 6}$$

This creates the final point estimate for the value of interest, μ, to be:

$$\overline{Y_a^*}(n) = \bar{Y}(n) - \hat{a}^*(n)(\bar{C} - v) = \bar{Y}(n) - \frac{\hat{C}_{cy}}{S_C^2(n)}(\bar{C} - v) \quad \text{(Law, 2007)} \qquad \text{Equation 7}$$

This is not the only form of estimating a*, although it is one of the more popular. The disadvantage of this method is that a* is no longer independent of C because it was created based on the simulation values of C (Law, 2007). This may result in some bias to the estimate (Law, 2007). The other methods for solving for a* include jackknifing, scenarios where we know the variance of C, or splitting up the simulation output data to estimate a* (Law, 2007). Glynn and

Szechtman (2001) state that according to Limit Theory, all estimates for a* are an improvement at the first-order central limit theory level and can only make a difference at the second-order level (Glynn & Szechtman, 2001).

2.2.3 Multiple Control Variates

Thus far we have examined the use of a single control variate. However, with simulations becoming even more extensive there is still the possibility of multiple control variates. They work the same way as the single case. Each CV can have its own a* associated with it depending on the correlation between the control variate and the resulting parameter. Because each CV will have an expectation of zero, with exception of bias introduced by a*, it can become even more helpful towards estimation and variance reduction as long as that bias is small enough to be accepted. This combination of CVs creates the following estimate for μ, using the same methodology as above and a total number of *k* control variates being applied,

$$\overline{Y_a^*}(n) = \overline{Y}(n) - \sum_{i=1}^{k} \hat{a}^*{}_i(n)(\overline{C}_i - v_i) \quad \text{(Law, 2007)} \qquad \text{Equation 8}$$

Now that μ has been estimated, the variance of the estimation is of interest as well. Since the equation introduces multiple control variates, it no longer benefits from just the correlation between a CV and response Y, but also the correlation between the CVs themselves and results in the following solution for variance (Law 2007):

$$\mathrm{Var}(Y_a) = \mathrm{Var}(Y) + \sum_{i=1}^{k} a_i^2 \,\mathrm{Var}(C_i) - 2\sum_{i=1}^{k} a_i \,\mathrm{Cov}(Y, C_i) + 2\sum_{j=2}^{k}\sum_{i=1}^{j-1} a_j a_i \,\mathrm{Cov}(C_j, C_i)$$

Equation 9

Derivating this equation with respect to each a_i will leave a set of k linear equations to solve for the variance minimizing weights for each a_i (Law 2007). Estimating these weights will lead to the same coefficients as a least-squares regression method, therefore when multiple CVs are used, it can be referred to as regression sampling (Law 2007). This is a key solution as it is the formula which will be used in solving for the values of each a* in the models in this paper.

$$\beta_{Coef} = (X'X)^{-1}X'Y$$

Equation 10

β_{Coef} = Coefficient on factors used to estimate the response
X = n by p matrix of design parameters used to get the response
Y = n by 1 matrix of responses

2.2.4 Sources for Control Variates

Averill Law discusses three different types of determining sources for control variates, as a well chosen control variate is highly important in this process. Kwon and Tew (1994) use two control variates, one highly correlated and one less correlated, and in every combination of runs, both control variates reduced the variance by some margin. This shows that while the degree of correlation does correspond to the impact the CV has, even less correlated variates can be helpful. Meanwhile, the most effective control variate is one that is highly correlated with the value of interest while having a low variance itself (Law, 2007). The first type is called an internal CV. This may be the most common and was seen in the earlier example of the service and interarrival times. They are the input random variates within a simulation (Law, 2007). Internal variates are created in order for the model to run (Law, 2007). This makes their application nearly free of cost making them worthwhile even if there is only a small change in variance (Law, 2007). While it may be tempting, it is not always best to use the same control variate for every response because the control variate may not be equally correlated with each response (Nelson & Yang, 1992).

Another source for this first example of internal control variates can also come from probability distributions. In many simulations, there is a node where an activity happens to only a percentage of the total entities that travel through it. The basic application of other internal control variates is the same. However, in this case we will standardize the values by their mean and their standard deviation using a slightly different equation used by (Bauer K. W., 1993).

$$CV_i = \sum\nolimits_1^N \frac{L_n - p}{\sqrt{Np(1-p)}} \qquad \text{Equation 11}$$

CV_i = Value for control variate i
L = 1 if entity n was true; 0 if entity n was false
p = Known probability of total entities that should be true
N =Total number of entities to pass through the node

Another type of CV is the external control variate. This type of CV uses a simplified version of the model to compute the expectation of the model's output random variable. This simplified model is created with assumptions that we may not be comfortable applying to the large model, but for the case of creating an external CV, can be very helpful (Law, 2007). The two models are then run simultaneously using CRN (Law, 2007). However, because this method creates a simplified model, it is not free of cost or time and these factors must be evaluated when deciding whether or not to apply the method.

The final type of CV, according to Law, would be the multiple estimators. These estimators are created when there are a collection of unbiased estimators for μ within the model. This collection is then compared to each other in the following format (Law, 2007):

$$\mathrm{Y_c} = \mathrm{Y}^{(1)} - \sum\nolimits_{\mathrm{i}=2}^{\mathrm{k}} \mathrm{b_i} \left(\mathrm{Y}^{(1)} - \mathrm{Y}^{(\mathrm{i})}\right) \qquad \text{Equation 12}$$

This method assumes that each Y^i for $i = 1 \ldots k$ is an unbiased estimator of μ so this method must be used carefully.

2.2.5 Problems with Control Variates

As control variates and all of these methods offer benefits to simulation experiments, they should not be used without thought. The simulation team cannot simply apply every variable within an experiment as a control variate (Law, 2007). As stated earlier, the bias that can be created by estimating a* can compound on itself as variates are added to the model and could quickly get out of hand if they are all chosen with no supporting correlation behind their selection. There are many methods for choosing the best a* and control variates in the model, one of which is shown by Bauer and Wilson (1992). However, they will not be investigated within the scope of this paper.

2.2.6 Where Control Variates have been Used

Control variates can be used in a large number of scenarios, and even applied to any stochastic simulation (Nelson B. L., 1990). Porta Nova and Wilson (1989) apply the method to a queuing network model. Kwon and Tew (1994) show an example for mechanics and technicians and a series of tasks. Henderson and Kim (2004) apply CVs to a discrete time-finite state space markov chain. Adewunmi and Aickelin (2012) describe its use for manufacturing, call center, and cross-docking discrete event simulations. Anonuevo and Nelson (1988) use CV for a model of an M/M/1 traffic system and a machine repair simulation. Nelson and Staum (1995) describe its use within financial engineering and ranking and selection systems as well as an inventory planning example. Nelson (1990) describe its use in predicting univariate mean, multivariate mean, and linear models before going into further detail with an example in machine repair, an inventory system, and an M/M/1 queue. Fort and Moulines (2008) apply CVs to financial scenarios including a call-put parity and asian option examples.

2.2.7 Other Variance Reduction Options

Control Variates are certainly not the only variance reduction technique available. Although it has its advantages, there are other possibilities that can be used effectively as well. Common random numbers (CRN) are the most commonly used reduction technique and take advantage of the same random number stream applied to alternative systems to evaluate their output (Adewunmi & Aickelin, 2012). Their use can be investigated further in Law (2007), Adewunmi and Aickelin (2012), and Nelson and Staum (1995). They are often combined with CVs so multiple systems can be compared easily because of the CRN, and with a smaller variance because of the CVs. Although it may assist in evaluation, Adewunmi and Aickelin (2012) found control variates to be the only method that was helpful in all three scenarios they investigated.

Antithetic variates are also a commonly used for variance reduction. Antithetic variates use random number streams that create a correlation between replications of the simulation model. This model is typically trying to improve the performance of a single system (Adewunmi & Aickelin, 2012). Of the three scenarios that were evaluated by Adewunmi and Aickelin they found antithetic variates to be particularly good for the cross docking scenario. Kwon and Tew (1994) combine CV and antithetic variates to investigate a potential improvement on prior systems. They concluded that all combinations reduced variance from a model which did not include any variance reduction techniques. However, the degree of improvement depended on amount of correlation between the control variates and the response (Kwon & Tew, 1994). The best model tested used antithetic variates for all random number streams as well as two control variates. More information on antithetic variates is covered by Law (2007), and the connection between CV and antithetic variates can be found in Glynn and Szechtman (2001).

The assumptions that allow control variates to be so useful are fairly simple. They include that the joint distribution for the control variates and the response are independently and identically distributed, that the expected value of the control variate is known, and that the variance of the control variate and the response are less than infinity (Fort & Moulines, 2008). These assumptions are covered in more detail in Nelson (1990) as they are needed for several theorems surrounding control variates and how to remedy the situation when the assumption is invalid or unknown. More information on the assumptions can also be found in Porta Nova and Wilson (1989), Glynn and Szechtman (2001), Nelson and Yang (1992), and Anonuevo and Nelson (1988).

Batching is a method that is used frequently in conjunction with control variates. It can be used to solve the problem of the absence of applying CVs for single replication experiment design of steady state simulation (Nelson 1990). It is also commonly used to remedy the bias that is introduced by a simulation that does not meet the normality assumption (Nelson 1990, Nelson and Yang 1992, Anonuevo and Nelson 1988). Jacknifing, splitting, and bootstrapping can be used to account for a lack of normality as well (Nelson B. L., 1990). Table 3 shows the areas of variance reduction that are covered by several different published journal articles.

2.3 Metamodeling

Control variates are typically implemented for their use in variance reduction of a single population simulation because of their low cost and ease of use. The control variates can also be used in multi population experiments whose output is used to create linear regression metamodels. These models show which factors are most important to the response and the coefficients associated with each factor. These metamodels can be used in place of the simulation to predict the response with some associated variance and confidence interval on both the response and the weights of the input factor levels. In this situation, control variates can sometimes be used to reduce the half width of the regression coefficient confidence intervals and to reduce the variance of the model. According to Arnold, Nozari, and Pegden (1984), when the variance is known, the addition of control variates will always be beneficial, while an unknown

Table 3: Guide for sources of prior research.							
Author	Subject Matter						
	Uses for CV	Derive CV	Batch Means	Antithetic Variates	Common Random Numbers	Description of Assumptions	Other VRT
(Porta Nova & Wilson, Nov. 1989)	X	X				X	
(Goodman, 2005)		X					X
(Kwon & Tew, 1994)	X	X		X			
(Glynn & Szechtman, 2001)		X		X		X	X
(Kim & Henderson, 2004)	X	X					
(Nelson & Yang, 1992)		X	X			X	
(Adewunmi & Aickelin, 2012)	X			X	X		
(Anonuevo & Nelson, 1988)	X	X	X			X	
(Nelson & Staum, 1995)	X	X			X		
(Nelson B. L., 1990)	X	X	X			X	
(Fort & Moulines, 2008)	X	X				X	
(Law, 2007)		X	X	X	X	X	X

variance opens up potential for additional control variates to be of little benefit. This shows why the selection of which control variates to use is important. Because the variance is rarely known, we will look into the unknown variance case in this thesis.

2.3.1 Linear Modeling

A linear model starts out in the form:

$$Y = \beta_0 + \beta_1 X_1 \ \ldots \beta_p X_p \qquad \text{Equation 13}$$

Where β are unknown, constant coefficients for factors X_i, i = 1…n and ε is the random error component. However, when control variates are added the form for the prediction equation must be modified to account for the added inputs. Now the model reflects the following form:

$$Y = \beta_0 + \beta_1 X_1 \ \dots \beta_p X_p + \delta'(C - \mu_c) \qquad \text{Equation 14}$$

Where there is now the added vector of coefficients δ, also represented as *a* in the univariate case for control variates explained earlier in the literature review. Each entry applies to the coefficient for a control variate in the vector C. The vector C is centered by subtracting each control variate by its corresponding distribution mean in the vector μ_c. For the remainder of this section, C will represent the centered value of $C - \mu_c$. However, because these control variates still have an expected value of 0, due to the known mean being subtracted, they should not change the expected value of the response making $Y = \beta_0 + \beta_1 X_1 \ \dots \beta_p X_p$ and $Y = \beta_0 + \beta_1 X_1 \ \dots \beta_n X_n + \delta'(C - \mu_c)$ equal.

Because we assume the joint normality of the responses and control variates, according to Arnold, Nozari, and Pegden (1984), we can create the conditional distribution as follows:

$$y_i | C_i \sim N_1(\mu_i + C_i' \textstyle\sum_C^{-1} \sum_{CY}, \sigma^2 - \sum_{YC} \sum_C^{-1} \sum_{CY}), \qquad \text{Equation 15}$$
$$Y \sim N_n(G\gamma, \tau^2 I_n) \text{ conditioned on } C$$

Where
y_i=response of observation i
C_i=vector of control variates of observation i
X_i=vector of factors of observation i
μ_i=mean for observation i
$\sum_{CY}$= covariance between control variates and responses
$\sum_C$= variance of the control variates
$G=(X_i \quad C_i)$
$\gamma = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_p \\ \alpha_1 \\ \vdots \\ \alpha_q \end{pmatrix}$
$\alpha = \sum_C^{-1} \sum_{CY}$
τ^2 = variance when conditioned on control variates

Because the variance is estimated using control variates in Equation 15, degrees of freedom to estimate the variance are lost due to being used for the control variate inclusion. This means the variance reduction achieved must now be greater to overcome the loss of degrees of freedom in the model error. Therefore, $\frac{n-p-q-1}{n-p-1}$ is the greatest value that $\frac{\tau^2}{\sigma^2}$ can take on in order to see benefits. This ratio, which varies based on the CV application, is referred to as the loss factor by Porta Nova and Wilson (Nov. 1989) and is used significantly throughout the research.

Arnold, Nozari, and Pegden (1984) then use this distribution formula to define $\hat{\gamma}$ and $\hat{\tau}^2$, the sample values for coefficients and variance respectively. The value of $\hat{\tau}^2$ and $\hat{\gamma}$ is shown in Equation 20 and Equation 21

$$\hat{\gamma}\,|C \sim N_{p+q}(\gamma\,, \tau^2(G'G)^{-1}) \qquad \text{Equation 16}$$

$$\hat{\tau}^2\,|C \sim (n-p-q)^{-1}\tau^2\chi^2{}_{n-p-q} \qquad \text{Equation 17}$$

$\hat{\gamma}\,|C$ and $\hat{\tau}^2\,|C$ are independent.

Where,

n=number of runs

p=number of factors

q=number of control variates

Arnold, Nozari, and Pegden (1984) then creates an estimate for the new coefficients of the factor levels to be:

$$\hat{\beta}_{CV} = [I_p\ 0]\hat{\gamma} \qquad \text{Equation 18}$$

$$\hat{\beta}_{CV}|C \sim N_p(\beta, \tau^2[(G'G)^{-1}]_{pp}) \qquad \text{Equation 19}$$

Where,

$\hat{\beta}_{CV}$= coefficients with use of control variates

$[(G'G)^{-1}]_{pp}$= the p x p upper left corner submatrix of $(G'G)^{-1}$

Because the expected value of the new $\hat{\beta}_{CV}$ is still equal to the true value β, it can be considered an unbiased estimator. For details on the model coefficient and variance derivations they can be found in Arnold, Nozari, and Pegden (1984). As mentioned in the control variate section, least squares regression analysis will solve for these coefficient values. With τ^2 being the variance estimate of $\hat{\beta}_{CV}$, it must be correctly estimated in order for accurate comparisons between models. This comparison will be used to evaluate the results of this thesis.

$$\hat{\tau}^2 = [(\bar{Y} - G\hat{\gamma})'(\bar{Y} - G\hat{\gamma})]/(n - (p + q)) \qquad \text{Equation 20}$$

$$\hat{\gamma} = (G'G)^{-1}G'Y \qquad \text{Equation 21}$$

This estimate can then be multiplied by $\frac{n-p-1}{n-p-q-1}$ in order to get a variance estimate that accounts for the loss of error degrees of freedom due to the number of control variates applied (Arnold, Nozari, & Pegden, 1984). These measures will be used to analyze the results.

Testing these new $\hat{\beta}_{CV}$ for statistical significance can also be accomplished where the null hypothesis sets Aβ=0 and an alternative hypothesis of Aβ≠0, for (p - k)p known matrix A of rank (p-k). According to Arnold, Nozari, and Pegden (1984), the result of the ordinary linear models creates a test statistic, f, where:

$$f = \frac{\hat{\beta}'_{CV}A'\left\{A[(G'G)^{-1}]_{pp}A'\right\}^{-1}A\hat{\beta}_{CV}}{(p-k)\hat{\tau}^2} \sim F_{p-k,n-(p+q)} \qquad \text{Equation 22}$$

$$\text{Prob}\left(f \le F^{\alpha}_{p-k,n-p-q}\right) = 1 - \alpha$$

$$\text{Reject if: } f > F^{\alpha}_{p-k,n-p-q}$$

Because these statements do not include the control variates, C, they are true unconditionally.

The Bonferroni approach can be used to create simultaneous confidence intervals of the coefficients. Montgomery (2009) mentions this technique is much simpler with very small reduction in accuracy when only a small number of coefficients are being estimated. This approach divides the alpha value of each estimation by the total number of parameters to be estimated concurrently. This way the simultaneous confidence interval is at least as large as the desired alpha value. The equation for confidence intervals with this method is:

$$\widehat{\beta_j} \pm t_{\frac{\alpha}{2p}, \; n-p-q-1} \sqrt{\frac{n-p-1}{n-p-q-1} \hat{\tau}^2 (X'X)^{-1}{}_{jj}} \qquad \text{Equation 23}$$

2.3.2 Multivariate Response

Porta Nova and Wilson (1989) took the work done with univariate responses and expanded it to multivariate responses. Now the linear model used to predict the response must account for all responses and takes on the form:

$$Y_{i.} = X_{i.}\beta + C_{i.}\delta + \varepsilon_i \;, \qquad 1 \le i \le n \qquad \text{Equation 24}$$

Where,
$Y_{i.}$ = 1 x m matrix, for run number, i, and each response variable up to m.
$X_{i.}$ = 1 x p matrix, for run number, i, and each factor of interest up to p.
β = m x n matrix, the coefficient for each factor, m, and design point n, within the design
$C_{i.}$ = 1 x q matrix, the control value for run number, i, and each factor of interest up to q.
δ = m x n matrix, the coefficient for each control variate, q, and design point n, within the design.
ε_i = 1 x m matrix, the residual error for each response up to m.

Through inspection it is seen how the univariate response model is expanded to account for multiple responses. Porta Nova and Wilson (1989) create many of the same derivations listed in the earlier univariate case, while this time accounting for multiple responses and the adjustments that must be made to the matrices throughout the process in order to accomplish them concurrently. They also apply the assumption for the joint normal distribution. The largest difference with additional responses is accounting for simultaneous confidence interval. Porta Nova and Wilson (1989) and Montgomery (2006) show how a simultaneous ellipsoid method, as well as the Bonferroni rectangle inequality, can be used to get concurrent confidence intervals on the factor coefficients by the use of F and student-t distributions respectively.

2.4 Measures of Effectiveness

2.4.1 Difference in Variance of Coefficients

Between Arnold, Nozari, and Pegden (1984), Porta Nova and Wilson (1989), and the other researchers mentioned previously, many measures of effectiveness were investigated to capture the impact that the application of control variates had on their respective examples. Similar measures will be applied later in the thesis to measure the effectiveness of the control variates when combined with different optimal design matrices. The most direct measure used by Arnold, Nozari, and Pegden (1984) is the difference of the variance values. If the $\mathrm{Var}(\hat{\beta}_{CV})$ is less than the $\mathrm{Var}(\beta)$, then the benefit of control variates is easily seen. Nozari et al derives the formula for $\mathrm{Var}(\hat{\beta}_{CV})$ and $\mathrm{Var}(\beta)$ to be the following:

$$\mathrm{Var}(\hat{\beta}_{CV}) = \frac{n-p-1}{n-p-q-1}\tau^2(X'X)^{-1} \text{ if } n-p-q>0 \qquad \text{Equation 25}$$

$$\mathrm{Var}(\beta) = \sigma^2(X'X)^{-1} \qquad \text{Equation 26}$$

$$\Delta = \mathrm{Var}(\beta) - \mathrm{Var}(\hat{\beta}_{CV}) = \sigma^2(X'X)^{-1} - \frac{n-p-1}{n-p-q-1}\tau^2(X'X)^{-1} = \left(\sigma^2 - \tau^2\frac{n-p-1}{n-p-q-1}\right)(X'X)^{-1} \qquad \text{Equation 27}$$

Because Δ needs to be greater than 0 to show improvement and $\frac{n-p-1}{n-p-q-1}$, the loss factor, will be greater than 1, there is potential for the $\mathrm{Var}(\hat{\beta}_{CV})$ to be greater than $\mathrm{Var}(\beta)$. However, as n increases towards infinity this value approaches 1 and therefore asymptotically it is better to apply control variates. Since an analyst usually does not have the resources to get a run size high enough to see this asymptotic effect, the selection of control variates is important so that the benefit of variance reduction is not counteracted by the loss of degrees of freedom when they are reallocated from estimating error to estimating the control variate values. Arnold, Nozari, and

Pegden (1984) describe that because $\frac{n-p-q-1}{n-p-1}$ is the largest that $\frac{\tau^2}{\sigma^2}$ can be while still seeing benefit of control variates, the analyst can create a target for variance reduction by getting this estimate and ensuring that the variance reduction be greater than that. This value will be shown in several tables in the results section to show the required variance reduction that must occur for control variates to be beneficial. If $\frac{\tau^2}{\sigma^2}$ is not less than the required ratio, multiplying by the loss factor will create variance estimates larger than they would be without control variates.

2.4.2 Expected Value of Square of Half-Width of Simultaneous Confidence Intervals

The next measure of effectiveness used by Arnold, Nozari, and Pegden (1984), compares the expected value of the square of the half length of the simultaneous confidence intervals. Therefore, we can compare the effectiveness of the control variate by dividing the expected square of half length using the control variate by the same value when the control variate is ignored. Arnold, Nozari, and Pegden (1984) apply the theorem that $E\left([(G'G)^{-1}]_{pp}\right) = \frac{n-p-1}{n-p-q-1}(X'X)^{-1}$ to assist in this transformation and get the simplified value of:

$$\left(\frac{\tau^2}{\sigma^2}\right)\left(\frac{F^{\alpha}_{p-k,n-p-q}}{F^{\alpha}_{p-k,n-p}}\right)\left(\frac{n-p-1}{n-p-q-1}\right); \text{where } \frac{\tau^2}{\sigma^2} < 1, \quad \text{Equation 28}$$

$$\frac{F^{\alpha}_{p-k,n-p-q}}{F^{\alpha}_{p-k,n-p}} > 1, \text{and } \frac{n-p-1}{n-p-q-1} > 1$$

This measure of effectiveness also illustrates how applying control variates under certain circumstances may not be beneficial. But again, as n grows to infinity, $\frac{F^{\alpha}_{p-k,n-p-q}}{F^{\alpha}_{p-k,n-p}}$ and $\frac{n-p-1}{n-p-q-1} = 1$. Arnold, Nozari, and Pegden (1984) also describe an upper bound for this technique because now $\frac{\tau^2}{\sigma^2}$ must be less than $\left(\frac{F^{\alpha}_{p-k,n-p}}{F^{\alpha}_{p-k,n-p-q}}\right)\left(\frac{n-p-q-1}{n-p-1}\right)$. Because of this relationship, Arnold, Nozari, and Pegden (1984) suggest that their "results indicate that, from the set of all possible control variates, a maximum number of $n - p - 2$ control variates should be used." However, because some of these equations were developed for only a single response, they may need to be modified when testing for simultaneous confidence intervals on multiple responses. Our experiments have a single response and we can therefore apply these measures with confidence. For more information on these techniques as well as the simple example on how additional control variates may not provide a better variance because of this relationship, please refer to Arnold, Nozari, and Pegden (1984).

2.4.3 Variance Ratio and Loss Factor

Porta Nova and Wilson (1989) expand the work done by Arnold, Nozari and Pegden (1984) to address conditions for multipopulation experiments with multiple responses. However, these require multiple replications for each design point. This replication may be part of the original experiment and have no impact on the overall cost and time of the experiment, while other scenarios do not have this replication built into the design due to constraints on cost, time, or resources, making this technique difficult. The first method is called the minimum variance ratio and was developed by Lavenberg et al (1982) then later used by Rubinstein and Marcus (1984) and Porta Nova and Wilson (1989). This compares the variance of the coefficients with the control variates to the variance of the coefficients without the control variates. If there is an improvement then the minimum variance ratio should be less than 1. Porta Nova and Wilson derive the equation for minimum variance ratio to be:

$$\eta(\Delta) \equiv \frac{|\mathrm{Cov}[\mathrm{vec}\beta(\delta)]|}{|\mathrm{Cov}[\mathrm{vec}\beta]|} = \left|I_m - {\sum_Y}^{-1}\sum_{YC}\sum_C^{-1}\sum_{CY}\right|^p = \left[\prod_{j=1}^{v}\left(1 - m_j^2\right)\right]^p \qquad \text{Equation 29}$$

Where,
v = min(m,q)
$m_1^2 \geq m_2^2 \geq \cdots > m_v^2 \geq 0$ are the ordered eigenvalues of ${\sum_Y}^{-1}\sum_{YC}\sum_C^{-1}\sum_{CY}$
m_j: $1 \leq j \leq v$ are the canonical correlations between Y and C

However, because we must estimate these parameters, the minimum variance ratio is adjusted by the loss factor, $\lambda(\hat{\Delta})$, to get the minimum variance ratio of the estimator, $\eta(\hat{\Delta}) = \lambda(\hat{\Delta})\eta(\Delta)$. Porta Nova derives the loss factor to be:

$$\lambda(\hat{\Delta}) = \left(\frac{n - p - 1}{n - p - q - 1}\right)^{mp} \qquad \text{Equation 30}$$

n = number of runs per replication
m = number of replications
p = number of factors
q = number of control variates

Similar to previous methods, if the number of runs is large enough, n will approach infinity and the loss factor reduces to one. But when the number of runs is not extremely large, the inclusion of control variates will force the loss factor to be less than 1.

Porta Nova and Wilson then extend the minimum variance ratio formula in Equation 29 to the predicted variance ratio:

$$\ddot{\eta}(\hat{\Delta}) = \hat{\eta}(\Delta)\lambda(\hat{\Delta}) \qquad \text{Equation 31}$$

In the predicted variance ratio formula, $\hat{\eta}(\Delta)$ is the estimated minimum variance and $\lambda(\hat{\Delta})$ is again the loss factor of the simulation. The estimated minimum variance uses the same formula as the minimum variance ratio but uses the pooled estimator values for all variances to create the following formula:

$$\hat{\eta}(\Delta) = \left| I_m - \hat{\Sigma}_Y^{\ -1} \hat{\Sigma}_{YC} \hat{\Sigma}_C^{-1} \hat{\Sigma}_{CY} \right|^p \qquad \text{Equation 32}$$

The predicted variance ratio there will hopefully be an accurate display of the improvement the control variates had on the simulation. This technique is an efficient way to measure the variance of the estimates, but due to the requirement of additional replications, could be too costly for many scenarios.

2.4.4 Confidence Interval

Confidence intervals for cases where more than one item is estimated at a time requires an adjustment to the confidence interval equation. Porta Nova and Wilson (Nov. 1989) introduce the Bonferroni inequality to show that multiple confidence regions must each have an individual α = (total α)/*m* for *m* concurrent confidence intervals, in the case of Porta Nova and Wilson *m* represents the number of different responses. However, as shown in Equation 23 and Montgomery (2009), *m* can also represent the total number of estimated half widths. In this research, all confidence intervals for the coefficients are two-sided and therefore *m* is equal to 2*p* where *p* is the number of factors in the model. Montgomery (2009) mentions the Bonferroni approach to be less accurate than the more complex oval method, but for the limited number of factors in this research it will be sufficient.

2.4.5 Selection of Control Variates for Inclusion in Model

As shown, selection of control variates is important towards seeing their optimal benefit. Nozari et al suggest two methods for finding the optimal combination of control variates that should be used in the actual prediction model. Because their implementation requires no additional runs, this analysis costs little to the analyst in terms of time and money. The first method is to construct the measure of efficacy for all combinations of control variates, while limiting the total number to n – p – 1. This is an exhaustive method which will take time but will ensure the use of the best combination. The second method is a forward selection method. Because it is an algorithm it may reduce the computation time, but there is no guarantee that the optimal combination will be found. In this method, the measure of efficacy will be computed for each control variate individually. The best variate will be selected and the measure of efficacy will be computed for all pairs that include the first variate. The best pair is then chosen and the step is repeated with all combinations of three control variates that include the previously chosen pair. This is continued until a combination of n – p – 2 control variates are found, until all combinations of variates are exhausted, or until no improvement is found. The final combination should be applied to the model. These are but two of the multitude of possible control variate selection methods possible. We suggest these because of their efficacy in the previous analysis with control variates done by Arnold, Nozari, and Pegden (1984).

2.4.6 Mean Square Error

Mean square error is a common way to measure prediction accuracy. It provides the average squared difference between the predicted value and the actual value. Squaring the difference between the values weights the differences so greater variations have a larger effect as well as projecting negative and positive differences onto the same plane. This requires a true value to be known which can be difficult in some scenarios. However, when creating a metamodel of a simulation, there are times when the simulation can be run a very large number of replications to remove variation and create an average response approaching the simulated true value. The following formula will be used for response predictions $\hat{y}_i$ on the true value y_i, across all *n* prediction points.

$$\text{Mean Square Error} = \left(\sum_{i=1}^{n}[(\hat{y}_i - \mathrm{y})^2]\right)/\mathrm{n} \qquad \text{Equation 33}$$

2.4.7 Prediction Half Width

Predicting a future observation can be a very important function of a metamodel. This allows the analyst to gain knowledge of the scenario without rerunning the simulation. This reduces vital time and resources and allows the analyst to make future predictions. Because this accounts for error of the prior observations as well as future prediction, this interval is greater than the confidence interval on currently observed points. The following equation is used to create the prediction half width which will be used during the analysis of this thesis.

$$\hat{y} \pm t_{\alpha/2,\ n-p-q-1}\sqrt{\frac{\mathrm{n-p-1}}{\mathrm{n-p-q-1}}\hat{\tau}^2(1 + (x_0(X'X)^{-1}x_0{}')} \qquad \text{Equation 34}$$

$\hat{y}$ = the point estimate from model
$\hat{\tau}^2$ = the variance estimate of the model
x_0 = the vector for the point of interest to be predicted
X = the design matrix for the model

2.5 Analysis of Covariance

A similar method to control variates is the implementation of analysis of covariance. A covariate is not the same as a control variate. This is implemented when a nuisance factor is uncontrollable. The factor is still measured on every run and then used to compensate its effect on the output variable (Montgomery D. , 2009). However it is typically applied to physical systems. Just like control variates, an adjustment is made on the observed output value in an attempt to get a more true response (Montgomery D. , 2009). However, in this case the mindset is to adjust the output so that the nuisance factor cannot inflate the response. Therefore, the significance test on the factors is more accurate as it removes potential interactions and influence from those nuisance factors. An appropriate statistical model involving a covariate would be:

$$y_{ij} = \mu + \tau_i + \beta(x_{ij} - \bar{x}_{..}) + \epsilon_{ij} \qquad \text{Equation 35}$$

where

y_{ij}= the *j* th response under the *i* th treatment

μ= overall mean

τ_i=effect of the ith treatment

β =the linear regression coefficient for the covariate

x_{ij}=the measurement of the covariate corresponding to y_{ij} (ij th run)

$\bar{x}_{..}$=the mean for all x_{ij} values

ϵ_{ij}=the error associated with the true value of y_{ij} and assumed to be normally distributed: (0,σ2)

It is easy to witness the similarities between this model and the model for control variates. As a run gets further from the mean, it has more of an impact on the output value. Similar to control variates, the implementation is fairly inexpensive and easy depending on the nuisance factor. For physical systems, it simply requires the monitoring of another factor. Depending on the factor this could be as easily as measuring the temperature that day to as difficult as measuring the instantaneous acceleration of particles. While this may be extensive in a few scenarios, there is usually the potential to monitor these extra areas with little effect on the overall test. In computer simulation scenarios this can be done at no cost in almost any case. By accounting for the effect of the covariate, we can more accurately judge the effect that the remaining factors have on the response.

2.5.1 Comparison to Control Variates

Although the analysis of covariance is very similar to control variates, they are different techniques that should be applied separately. This thesis has chosen to apply control variates because as we investigate metamodel simulations, the true mean values of the control variates will be known. In the analysis of covariance, the values for the covariates are compared to the sample mean. This sample mean may reflect the true mean, but due to the naturally occurring randomization of any situation, whether it is a computer simulation or real-world test, variance from the true mean occurring because of the randomization will keep the metamodel from equaling its true value. As described by Montgomery (2009), analysis of covariance has a place in design of experiments and modeling, however, that place is not when the true mean value is known, because then the true value can be used to create a more accurate model.

Because the analysis of covariance is based on concomitant variables that are compared to their sample mean, their coefficient values within the metamodel are used for future predictions. In comparison, the coefficient values for the covariate factors of control variates are removed from the metamodel. As discussed in the control variates section of this chapter, the expected value of the control variates is zero, making them non-biased estimators that can help create a metamodel with a smaller half-width on each coefficient if the control variate was indeed a helpful factor in reducing variance.

2.5.2 Assumptions

The assumptions for analysis of covariance models are the same as regression models and analysis of variance models (Quinn & Keough, 2001). The error terms from the fitted model found using analysis of covariance are assumed to be independently and normally distributed with similar variance between groups. By plotting residuals versus adjusted group means, the assumption of homogeneous variance can be easily checked for satisfaction of this requirement. Just like another regression model, nonhomogeneous variance can be corrected with transformations to the response. Other assumptions covered by Quinn and Keough (2001) include linearity between the factors and the response, covariate values are similar across groups, and the covariates are fixed variables.

2.5.3 Problems of Implementation

According to Miller and Chapman (2001), analysis of covariance is often misused because the reviewers have no means of achieving the goal of correcting or controlling for real group differences on the concomitant variable. Another common problem with using analysis of covariance is when there is an interaction between the concomitant variable and the independent factors. This correlation should be low in order for analysis of covariance to be an effective tool (Tabachnick & Fidell, 1996). Without a plan from the beginning, many of these techniques fail to make a significant difference. They say it is often used in psychopathology research, but if it is misused it simply takes away from its potential in further research and development. In this thesis, control variates are chosen over analysis of covariance because of the known mean values for the control variates. Not having a goal can affect the use of control variates as well. In this case, we know the techniques that will used in order to investigate impacts on simulation metamodels. Therefore, we apply an optimal design to create a metamodel of a simulation that includes control variates in order to reduce variance on the prediction values, as well as the variance on the coefficients within the metamodel.

2.6 Optimal Designs

2.6.1 Overview

A major part of creating an experiment or simulation is deciding on the structure of design the experiment will use. As mentioned previously, the variance for the coefficients is found with the equation $\sigma^2(X'X)^{-1}$. While the control variates have an effect on reducing the value for σ^2, the experimental designs will affect the output of $(X'X)^{-1}$ also influencing the coefficient and other outputs. The structure of the design will assist in finding the needed information in the most accurate way possible (Montgomery D. , 2009). For example, some experiments allow the analyst to check defined points of interest based on the contract or physical constraints. When this is not the case, the analyst may have a much larger range of potential settings they are interested in, and this area is where design of experiments can be applied in order to get the most benefit from the limited information (Law, 2007). When resources and costs are not restricted a space filling design may be constructed. However, in the case of this research and many real world scenarios, there is a limit to the number of allowable runs making space filling designs impractical and optimal designs very appealing. In this scenario, the test team decides what alternative configurations as well as how they plan on evaluating and comparing the results. For this purpose, there are several universally popular designs such as full factorial designs or Latin square designs that can be applied (Montgomery D. , 2009). However, when funds and/or resources are limited it is not always feasible to do a full factorial design. In addition, as the number of factors increase the number of needed replications can quickly get out of hand. When there are restrictions within the experimental region that make full factorial designs infeasible, computer generated designs may also be helpful. For example, a limit on the sum of two factors or when a nonstandard model is being investigated, such as a quartic model or a response surface problem with categorical variables (Montgomery D. , 2009) are cases that benefit from computer generated designs. Optimal design theory has also been used effectively in developing polynomial models over irregularly shaped regions, such as mixture design problems (Montgomery, Peck, & Vining, 2006). For all of these cases, it is still important for the test team to decide what they are attempting to find out from the test. They should then decide on a test criterion for which the design will be evaluated.

2.6.2 Properties of Optimal Designs

There are several different criterions on which designs can be evaluated. They are usually found with the help of a computer and therefore called computer-generated designs. For example, a design could be A-optimal, D-optimal, G-optimal, I-optimal, or V-optimal (Montgomery D. , 2009). Each of these criterions evaluate the design with a different metric and offer a different advantage to the test team. Some criteria will return the same optimal design as others, while others may create slightly different or even very different optimal designs depending on the criterion chosen (Montgomery D. , 2009). For example, the full factorial design is A, D, G, I, and V optimal for fitting the first-order model in k variables or with interaction, which is one reason it is so widely used in the testing world (Montgomery D. , 2009).

These optimal criteria are used to achieve certain properties in the moment matrix M:

$$M = \frac{X'X}{N}$$

Where X is as defined for Equation 10 and N is the total number of experimental design points

These elements are important in determining the rotatability of the design (Myers & Montgomery, 2002). When the $\mathrm{Var}[\hat{y}(x)]$ depends on the distance from the center of the design and not its direction, then it is called a rotatable response surface design (Montgomery, Peck, & Vining, 2006). This is an important property of optimal designs because when solving for the design, the orientation of the points is often times unknown, and the distance from center, regardless of direction is what is used (Montgomery, Peck, & Vining, 2006). This makes all points equally important as long as they are the same distance from center. If the design is not rotatable then the estimates could be very different at different points within the design region (Montgomery, Peck, & Vining, 2006).

2.6.3 A-Optimal Design

The A-optimal design criteria's focus is to find a design that minimizes the individual variance of the regression coefficients (Montgomery D. , 2009). The variance of regression coefficient β_i, i = 1… p is a function of the corresponding diagonal of the $(X'X)^{-1}$ matrix. If we let $H = (X'X)^{-1}$ then the variance of a coefficient β_i, $Var(\beta_i) = \sigma^2 H_{ii}$. Therefore, a design is considered A-optimal when it minimizes the sum of the main diagonal elements of $(X'X)^{-1}$, or the trace of $(X'X)^{-1}$ weighted by N (Montgomery D. , 2009). In relation to the moment matrix, M, this can be defined as the:

$$\mathrm{Min}(\mathrm{tr}[M(\zeta)]^{-1}) \qquad \text{Equation 36}$$

Where ζ is all potential designs (Myers & Montgomery, 2002). In other words, it minimizes the sum of variances of the regression coefficients (Montgomery D. , 2009). However, unlike the soon to be reviewed D-optimality, it does not account for the covariance among the coefficients. While A-optimal designs can be generated within some computer programs, it is not popular enough to be included in all software packages.

2.6.4 D-Optimal Design

D-optimal designs are typically the most common of the designs used that have an optimal design focus (Montgomery D. , 2009). A design is considered D-optimal when the determinant of $(X'X)^{-1}$ is minimized, and specified using *D* because of evaluating the determinant (Montgomery D. , 2009). Constructing a D-optimal design will minimize the volume of the joint confidence region on the vector of regression coefficients (Montgomery D. , 2009). A simple way to compare two designs for this criterion is the following formula:

$$D_e = \left(\frac{\left| \left(X'_2 X_2 \right)^{-1} \right|}{\left| \left(X'_1 X_1 \right)^{-1} \right|} \right)^{1/p}$$ Equation 37

Where X_1 and X_2 are the design matrices for the designs we would like to compare and p is the number of model parameters (Montgomery D. , 2009). D-optimality is great for variance-optimal designs in first-order and first-order with interaction models, however, as models get slightly more complex and moving to second-order designs, D-optimality can still be considered due to its simplicity (Myers & Montgomery, 2002). It has also been suggested that if additional runs are being conducted, the placement of these runs should move towards D-optimality (Montgomery, Peck, & Vining, 2006). One common problem with D-optimal designs is that they often consist of runs where the number of points is equal to the parameters, this makes model adequacy checking not possible (Montgomery, Peck, & Vining, 2006). Since it is the most widely used design for simple cases it can be easily constructed within popular software packages such as JMP, Design-Expert, and Minitab (Montgomery D. , 2009).

2.6.5 G-Optimal Design

While D-optimal designs may be the most widely used overall criterion, G-optimal criterion is the most popular when concerned with the prediction of the response and looking for a prediction variance criteria (Montgomery 2009). G-optimality occurs when the design minimizes the maximum scaled prediction variance, V(X), over the design region as follows:

$$\mathrm{Min}\ [\mathrm{Max}\ \mathrm{v(x)}];\ \mathrm{v(x)} = \frac{\mathrm{N} \cdot \mathrm{V}[\hat{\mathrm{y}}(\mathrm{x})]}{\sigma^2}\ \text{and}\ \mathrm{x} \in \mathrm{R}$$ Equation 38

where N is the number of points in the design (Montgomery 2009). Another metric when using G-optimality is the G-efficiency of a design which is evaluated with the formula (Montgomery 2009):

$$G_{eff} = \frac{p}{\mathrm{Max}\, v(x), x \in R} \qquad \text{Equation 39}$$

G-optimal designs are extremely common and because $\mathrm{Max}\, v(x), x \in R = p$ for two-level designs with a resolution greater than III and levels at ±1, the design will have an efficiency of 1 and be G-optimal for first order models (Myers & Montgomery, 2002).

2.6.6 I-Optimal Design

I-optimal designs can also be referred to as the integrated, IV, or Q criterion, and can be easily constructed in JMP and other popular statistics software (Montgomery D. , 2009). They measure the design with a single performance statistic which is found by averaging the scaled prediction variance over some region of interest (Myers & Montgomery, 2002). These designs are evaluated with the following formula (Montgomery D. , 2009):

$$I = \frac{1}{A}\int_{-1}^{1}\int_{-1}^{1} V[\hat{y}(x_1 x_2)]dx_1 dx_2 \qquad \text{Equation 40}$$

By integrating over the entire region and dividing by the area of the region, we receive the average for the entire area. This criterion is considered optimal with 2k designs for fitting first-order models as well as first order models that include interaction (Montgomery D. , 2009). One of the reasons for the popularity of the I-optimal design is that it is easy to conceptualize the average prediction variance over a region (Myers & Montgomery, 2002). Comparing designs is similar to other criterion with the use of an efficiency equation. I efficiency can be found for design ζ* as (Myers & Montgomery, 2002):

$$I_{eff} = \frac{[I(\zeta)]}{I(\zeta^*)}, \mathrm{Min}\, \zeta \qquad \text{Equation 41}$$

Again, the I-optimal criterion is optimal for two-level first-order orthogonal designs with resolution greater than III and levels of ±1 (Myers and Montgomery 2002).

2.6.7 V-Optimal Design

A more specific form of I-optimal designs are called V-optimal designs. These designs also focus on prediction variance but over a specific collection of points of interest rather than the entire design region (Montgomery D. , 2009). These points could be a candidate set by which the design was chosen, or simply a collection of points hand chosen by the test team because they find them specifically important (Montgomery D. , 2009). Regardless of the reason, the set of points are specifically chosen by the experimenter. The design is considered V-optimal when these pre-identified points have a minimum average prediction variance (Montgomery D. , 2009). Because of the complexity of solving for these values, computer software is used often.

2.6.8 Alias-Optimal Design

An alias optimal design is concerned with limiting the effects of aliased terms. One disadvantage of the standard optimal designs previously mentioned is that it does not consider the aliasing between specific model terms and those that may be important but not included in the model (Jones & Nachtsheim, 2011). The alias matrix compares the matrix of model effects, X_1, and the matrix of all aliased effects, X_2, to create an aliasing matrix.

$$A = [X'_1X_1]^{-1}X'_1X_2 \qquad \text{Equation 42}$$

The goal of the optimal design is to minimize the *tr(AA')*, or minimize the sum of the squared diagonal elements of *A.* This design criterion has the clear advantage of still being a better predictor if interactions are added later because there is much lower risk of them being aliased with another important factor. This allows the analyst to determine which factor or interaction is the true cause for changes to the response.

2.6.9 Problems with Optimal Designs

As mentioned in many of these criteria methods, standard designs are optimal for first order and first order with interaction models. However, that is not always the model of interest in an experiment. Now it is important to remember that optimal design criteria evaluate the design on one area, but the analyst may be interested in a model that can perform well in several measures of efficacy. Therefore, the simple first order model with a standard design may be easy to think of but can quickly become poor in practice even though it was effective in theory. Standard response surface methodology designs are rarely optimal for second order models but they are constructed to achieve many desirable properties such as simplicity and still near optimal solutions (Myers & Montgomery, 2002).

2.6.10 Solving for Optimal Design

The optimal criterion values are often very difficult to find, hence one reason non-optimal designs are sometimes used. For this reason, computer software, with help from an algorithm, must be used to discover it. According to Douglas Montgomery (2009), a point exchange algorithm is very common. This algorithm takes an initial set of points chosen by the test team, and then exchanges points for alternative options in order to find an improved design (Montgomery D. , 2009). Although not every possible design is checked, this algorithm will still find a solution very close to the true optimal design. The coordinate exchange algorithm is another option for constructing the optimal design as described by Montgomery (2009). This algorithm takes the initial design and searches over each coordinate of the points until no improvement can be found (Montgomery D. , 2009). The initial design, unlike the point exchange algorithm, is randomly selected and run through multiple times to improve the chances of getting the optimal design.

2.6.11 Use of Optimal Designs with Metamodels

Although these optimal designs, and their uses, are typically thought to be used on physical tests with limited funds, time, and/or resources, they can also be used in computer simulations. As computers become more powerful they are now being used to model even more complex systems which will still challenge the computing power of the computer itself. In these cases, minimizing the number of runs and making the most of the ones that occur can be very important when prediction power is important. Montgomery (2009) mentions factory planning, scheduling models, traffic flow simulators, and Monte Carlo simulations that sample probability distributions as examples of simulations that commonly employ optimal designs within their simulations to model a real-world process. Law (2007) mentions simulation based optimization being used to evaluate direct economic importance outputs such as profit or cost by testing all possible combinations of the input factors. He also lists emergency-room operations, automobile manufacturing, and management of a production-inventory system as potential applications for optimal simulations. A collection of authors list several other areas where optimal designs and simulation can be applied, such as, economic uncertainty and high speed civil transport vehicle (Bandte & Mavris, 1995); peace-keeping mission modeling (Johnson, Lampe, & Seichter, 2009); strength and accelerated life testing of a device (Chernoff, 1962); engineering and management science (Kleijnen J. P., 2005); combustion circuit design, controlled nuclear fusion device, plant ecology, and thermal energy storage (Mitchell, Sacks, Welch, & Wynn, 1989); and process or device design, simulator tuning, process control recipe generation, and statistical process or device design (Boning & Mozumder, 1994).

As these systems get more complicated, the simple full factorial designs for first-order models are not as realistic. This is when a computer generated optimal design can become even more helpful. Other helpful models that are used commonly throughout computer simulations include space-filling designs such as the latin hypercube design, and are popular because although they do not include replications, the design points are usually evenly spread throughout the design region (Montgomery D. , 2009). Uniform designs, Gaussian process models, and maximum entropy designs are also potential design options. Although, they will not be discussed in this paper, further information on these and other commonly used designs can be found in Montgomery (2009), Pronzato and Walter (1990), Mitchell et al (1989), and Boning and Mozumder (1994).

Using design of experiments within simulations is a growing field. This is what Douglas Montgomery (2009) had to say on the area:

> "These experiments with computer models represent a relatively new and challenging area for both researchers and practitioners in RSM and in the broader engineering community. The use of well-designed experiments with engineering computer models for product design is potentially a very effective way to enhance the productivity of the engineering design and development community"

In real world designs the test team must control factors or account for uncontrollable factors. The use of simulation allows the team to control everything and investigate the results. Even the variability can be accounted for by using random number generators, a simulation run can be replicated for additional information in the future (Law, 2007). While simulation allows for these uses, the additional application of design of experiments allows the test team to do so in the most efficient way possible, saving the organization time and money, while creating a better model. Since these models are constantly getting more complex, this savings and improvement is of gaining concern and software packages are being developed and used more frequently. These packages are then interfaced with simulation programs to get the most out of the operating system. Law (2007) lists AutoStat, Extend Optimizer, OptQuest, SimRunner2, and WITNESS Optimizer as optimization programs that incorporate simulation software and can be researched separately for their differences, advantages, and disadvantages. However, what they all have in common is using simulations, and an extensive amount of computing power, to get results the user is interested in investigating. With the addition of design of experiments and control variates to a general simulation, both the results and computing time can be improved as will be shown in this paper.

2.6.12 Potential Areas for Use of Control Variates

There are several examples in literature on the application of DOE simulations that could potentially benefited from the use of control variates. Bandte and Mavris (1995) mention the need for 1000 to 10000 runs for a good representation of the probability distribution when applied to engineering analysis of a high speed civil transport vehicle. Control variates within the simulation could potentially reduce the variance of the results enough to reduce the runs without sacrificing the representation potential. Baesler et al (2003), use DOE for estimating the maximum capacity in a Chilean emergency room. Their process uses a simulation model of the emergency room, and values chosen by a design matrix to predict the behavior of the patient's time in the system and the maximum possible demand that the system can handle. Although DOE was used to define the minimum number of physical and human resources required to serve these demands, the addition of control variates could potentially reduce the variance of the design even further. This could make the predictions more reliable by reducing the confidence interval and closing in on the true values. Johnson et al (2009), constructs a simulation of a refugee camp and uses a DOE test matrix to get potential outcomes from the simulation and limit the options for settings within the simulation. In this case, control variates could take the known values in the system to receive more accurate results from each setting so the test matrix can zero in on the best values for the simulation with more confidence.

These are hypothetical possibilities for the use of control variates and the only way to find out if they would actually help the system would be to implement them. However, since the cost of a control variate is free due to all the information already being known within the simulation, it is still something that can be investigating cheaply and easily. This is a tool that has yet to be used in these scenarios, or any of the listed areas where optimal designs can be applied. Although design of experiments and optimal design criteria has assisted in reducing variance and ensuring the results are exactly what the test team is concerned with, there is always the question of how to improve the test even further so that the results can have improved reliability with fewer replications and the additional use of control variates should be one step towards achieving this.

2.7 Summary

This literature review has covered the foundational material used in this thesis. The use of simulations to create metamodels for prediction purposes has been highlighted as they allow the analyst to further understand what is happening in a process, how things impact the process, and what may result from the process of interest. Optimal designs and control variates are two ways to increase the benefits of metamodeling even further by creating prediction models with lower variance and a better prediction. However, each step and technique must be fully understood and implemented correctly in order to achieve the necessary benefit. Control variates have been common in many areas to reduce prediction variance by comparing the relationship between a known value and the actual value within the simulation run. Optimal designs are often used to investigate the design space as efficiently as possible. If interested, Table 4 shows where to conduct further research on a specific area of optimal designs. The next chapter, methodology, shows more specifically how both methods will be implemented on two simulations of interest so that potential results can then be evaluated effectively.

Table 4: Table of sources for prior research in optimal designs.

Authors	Classifiers within Optimal Design Literature				
	Uses of DOE	Uses of Simulation	Potential for CV	Description of Designs	Description of Optimal Criteria
(Baesler, DaCosta, & Jahnsen, 2003)	X	X	X		
(Bandte & Mavris, 1995)	X	X	X		
(Boning & Mozumder, 1994)	X	X		X	
(Chernoff, 1962)	X				
(Johnson, Lampe, & Seichter, 2009)	X	X	X		
(Kelton, 2000)		X		X	
(Kleijnen J. P., 2005)		X			X
(Law, 2007)		X		X	X
(Mitchell, Sacks, Welch, & Wynn, 1989)	X	X			
(Montgomery D. , 2009)	X			X	X
(Myers & Montgomery, 2002)				X	X
(Montgomery, Peck, & Vining, 2006)					X
(Pronzato & Walter, 1990)				X	X
This Thesis	X	X	X	X	X

3. Methodology

3.1 Chapter Overview

This chapter covers the specific techniques used to combine control variates, optimal designs, and statistical analysis to more accurately and efficiently characterize a simulated process with a metamodel. Two simulation examples will be used to demonstrate the implementation of control variates. The first is an adaptation of an example used by Arnold, Nozari, and Pegden (1984) that examines the average waiting time of cars on a one-lane section of road. The second simulation was generated for the 618th TACC to examine issues surrounding the time and personnel requirements for flight planning. Each simulation has several potential factors, distributions, and control variates which may be used to create the optimal designs and models used to predict the responses of interest. Several statistical measures are used to compare the effectiveness of the control variates with each design structure. I hypothesize that control variates will provide reduced variation and prediction error across all optimal designs.

3.2 Description of Research Methodology or Approach

Variance is a natural occurring part of any process. Simulations are becoming a more common way to model these scenarios and help account for this variance so the analyst has improved knowledge on what is occurring in the process. The ability to reduce variance for metamodels generated from simulation output will improve the prediction ability, and increase the overall understanding of the scenario. Control variates and optimal designs are two ways to achieve this goal. This research combines these two methods and investigates their impacts.

As previously discussed, two simulations will be used to illustrate the benefits of employing these variance reduction techniques. Figure 1 illustrates graphically how the different parts come together. The response, factor, and control variates for each simulation will be determined based on the analytical goals of the simulation. Subsequently the experimental design region will be determined and optimal designs will be created to more efficiently characterize the simulation output over the entire design region. Space filling designs may work great for these scenarios as well, but with this research has an interest in restricting the number of available runs. This restriction makes optimal designs more appealing than space filling designs.

D, Alias, and I optimal designs used for the analysis will be constructed using JMP10. Provided the factors of interest, their constraints, number of runs, and the design criteria of interest; JMP returns a design optimal to the desired criteria. Although this is a difficult process by hand due to the complicated equations described in the methodology, JMP and other popular statistical programs can make it an easy process.

Constraints are created to restrict the design region and put an emphasis on each optimal criterion. They are input into JMP along with the factors when creating the different designs. Constraining the design space will ensure different designs for each optimal criterion and force the program away from full factorial or half fractional designs.

The designs will be run through the appropriate ARENA model using the ARENA Process Analyzer. The simulation output is then used to construct a second order metamodel. The designs and outputs are input in JMP to create the best possible model without considering control variates using the stepwise function. This is the process done by Arnold, Nozari, and Pegden (1984), Porta Nova and Wilson (Nov. 1989), and several others. However, due to the inclusion of control variates and the resulting reduction in variance, some additional factors may also become significant. Therefore, a new model will be created using the stepwise function to see if any additional factors are significant when control variates are included. This process will be done for all experimental designs and responses. The significant metamodels are then used to calculate the coefficient half widths and variance estimates discussed in the literature review with the MATLAB code shown in Appendix A.

These values are compared and speculation could be made about which models showed benefits from control variates and which did not. The next step will show whether that speculation can be supported when the models attempt to predict a number of random points. These points were randomly selected throughout the design region. Predicted mean square error is used to evaluate all models, while prediction half width and coverage will also be used to judge the efficacy of control variates. Again, comparing the performance of the models in the end, to what we knew about the model at the beginning allows us to develop recommendations on when control variates should be employed in real world scenarios and when their benefit may not be seen.

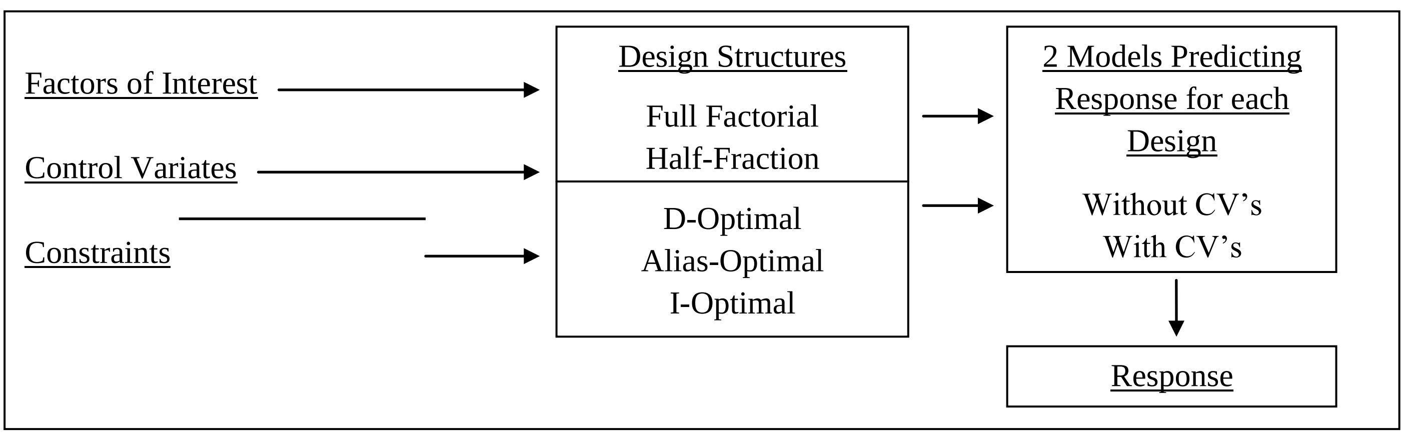


Figure 1: Graphical representation of the parts of the simulation analysis

Practical application of control variates will be highlighted by examining the model accuracy improvements using control variates to predict a random selection of points in the design region. Figure 1shows the requirements for constructing the models. Each design requires the factors, control variates, and constraints. This creates a model with and a model without control variates, each of which used to predict the response of interest. Figure 2 shows how each model will be used to gain information on the response, as well as the measures used to compare the model with control variates and the model without control variates.

3.3 Methodological Assumptions

Many assumptions are inherent to the techniques used. Control variates require the assumption of joint normality between the control variates and the response. They also assume that each design has constant variance. This latter assumption is not as widely accepted and will be investigated slightly in the research. These assumptions were discussed in more detail in the previous literature review section, under control variate assumptions. Optimal designs also have their own set of assumptions. This includes knowledge of the design space is complete and reflective of the actual situation to be examined. General assumptions for this research also includes that all models are valid and reflective of their scenario, prior work which this is based on is correct, and the software used creates accurate distributions and calculations.

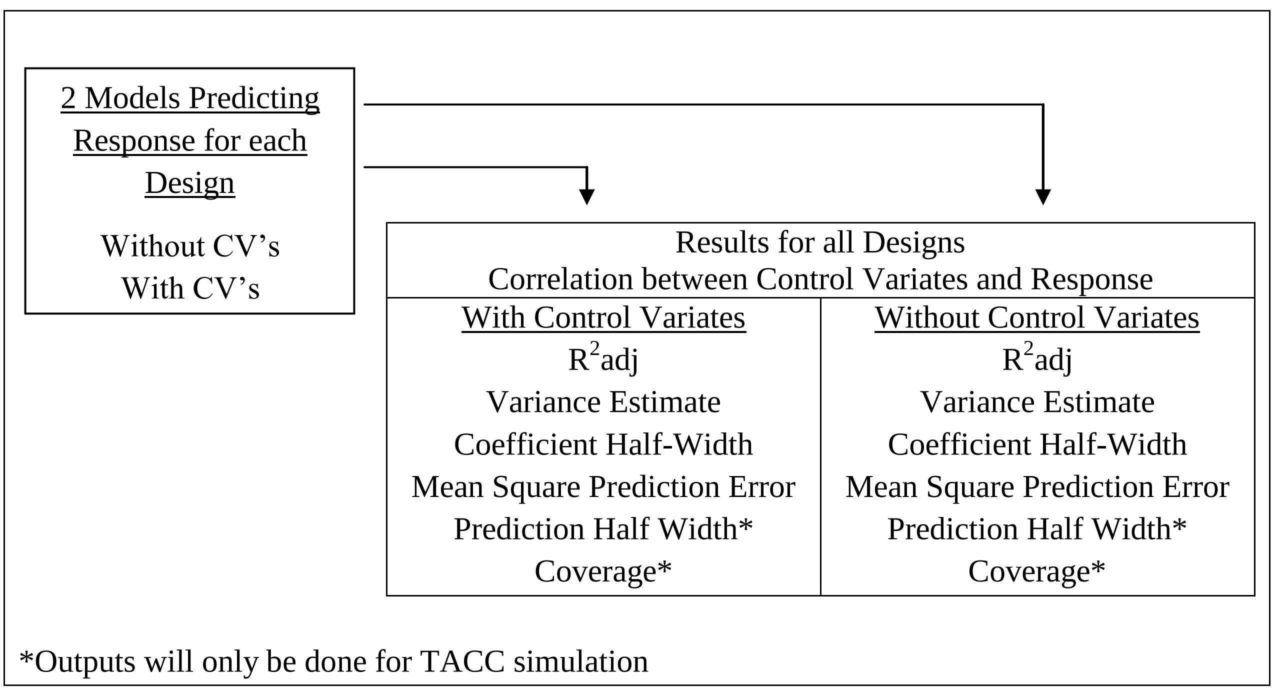


Figure 2: Graphical representation of the results for each design

3.4 Data

3.4.1 Inputs

The calculation of measures of effectiveness requires the design matrix, each control variate, and the responses as inputs in order to output the appropriate parameter measurements. The design matrix is a *k* by *p*+1 sized matrix where each design point is a separate row, and each column represents a factor or interaction to be included in the model as well as a column vector of ones to find the intercept. These will be the optimal designs found using JMP. Each control variate requires its own matrix. The response matrix is similar to a control variate matrix, where each entry represents the value from that run.

3.4.2 Results Analysis

Once the outputs for response and control variates are known, these values are used to calculate the statistical measures used to determine the efficacy of the designs. The statistical measures of effectiveness chosen include the variance estimate, coefficient values of the model, and the half width of the coefficients. For the TACC simulation, this also includes the half width and coverage of the prediction points. Equation 10, $\beta_{Coef} = (G'G)^{-1}G'Y$, is used to solve for the coefficients where G is the matrix for all designs points and the control variates of choice. Equation 23, $\widehat{\beta_J} \pm t_{\alpha/(2p),\ \mathrm{n-p-q-1}}\sqrt{\frac{\mathrm{n-p-1}}{\mathrm{n-p-q-1}}\hat{\tau}^2 (X'X)^{-1}{}_{jj}}$, provides simultaneous half widths for factor *j*. A smaller half width is obviously desirable, but also is a smaller variance estimate. Equation 20**Error! Reference source not found.**, $\hat{\tau}^2 = \frac{[(\bar{Y}-G\hat{\gamma})'(\bar{Y}-G\hat{\gamma})]}{(n-(p+q))}$, as well as the variance estimation equations involving degrees of freedom, $\hat{\tau}^2\,\frac{\mathrm{n-p-1}}{\mathrm{n-p-q-1}}$ are employed as measures of efficacy. Clearly, a smaller value is desired. Although control variates should reduce the variance estimate, the degrees of freedom lost in error estimation may do more damage when estimating the value than the help in variance reduction. This comparison will show whether the control variate is worthwhile across all four design structures. Calculation of these statistics is accomplished using MATLAB R 2012a. The script file to perform the calculations may be found in Appendix A.

Once these inputs are in place and the code is run, MATLAB will then output a matrix for the variance, the coefficients, and the half width estimates for a model with no control variates and each possible combination of control variates. These matrices can be easily compared to see which combination of control variates, if any, offer a reduction in variance or half width.

The coefficient values are then used as an input into Excel to create the prediction values of a number of randomly generated points within the design space. These points are found using Microsoft Excel's random function to randomly pull a number between -1 and 1 for each factor. These points must also satisfy the constraints which were used to create the optimal designs. We do not want to attempt to predict a point that was infeasible when creating the optimal designs. Mean square error when predicting these points are used as a measure of the prediction accuracy of each model. In order to find the "true" value for each random point, the model was run 1500 times at each random point. The average response of the 1500 runs is accepted as truth for that specific point in the design space. The models created will then be used to predict the random points. Equation 33, Mean Square Error = $[\sum_{i=1}^{n}(\hat{y}_i - \mathrm{y})^2]/n$, will be used for this calculation. Again, clearly a lower value for mean square error is more desirable as it shows the model is more accurate at predicting the true long term average at a design point. The TACC simulation will also examine the half width and coverage of the prediction points using Equation 34. Figure

3 gives a clear graphical representation of how the inputs and outputs are related with each software platform.

If required this process could be done again with new designs, new factors, or using these outputs to adjust the inputs in an additional iteration. Those decisions will be made once the initial results are found but by collecting the data in the above mentioned manner, it will be easily available for further adjustments.

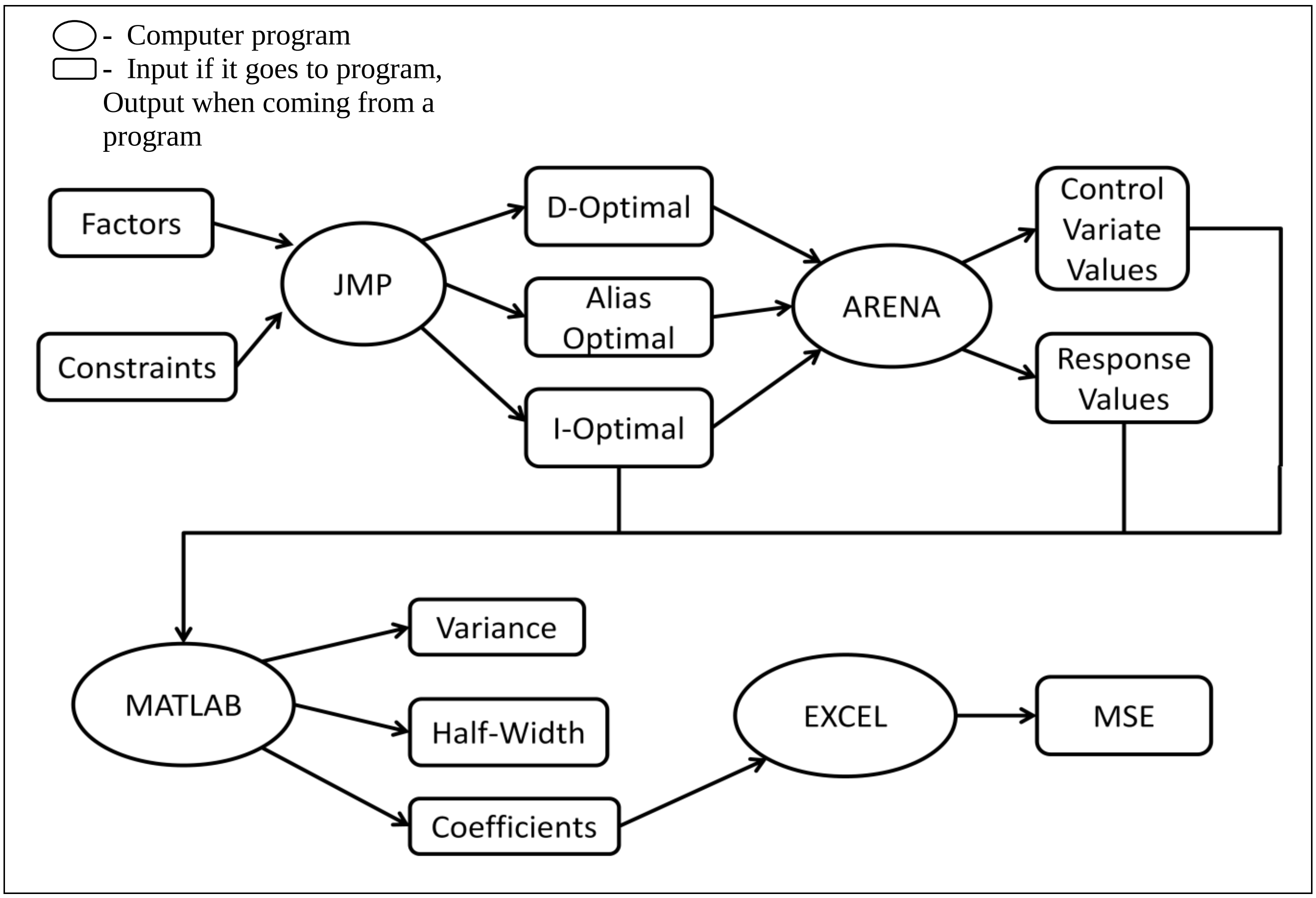


Figure 3: Image depicting the inputs and outputs through the different software packages used.

3.5 Limitations

Limitations of this research exist as we do not investigate all potential applications, variations of the techniques, or simulations possible with control variates and optimal experimental designs. Analytical assumptions made for this analysis have been stated throughout the methodology chapter. Insights gained are based on the use on only two simulations. Ideally these techniques would be combined and shown across a wide variety of scenarios. However, in this case I believe two different simulations are enough for initial investigation for potential benefits. The two simulations used will be fairly small compared to some extremely large scale simulations currently used. This will allow for easier manipulation for this research. Although it may not be as easy to see the benefit as simulations grow extremely large, the potential will be shown and the limited cost of implementation could prove it worthwhile.

3.6 Research Design

3.6.1 Traffic Lane Simulation

3.6.1.1 Simulation Description

The first simulation used in this research is a variation of the one used by Arnold, Nozari, and Pegden (1984). They title the example as "Single Lane Traffic Analysis." It simulates a 500 m section of road where the two-lane traffic flow, one lane in each direction, has been reduced to one lane due to construction or another barrier. There are traffic lights on each side where a green light obviously allows the cars to proceed down the single lane while a red light forces them to stop so the other direction can use the road, see Figure 4. Each side has a different interarrival time, both distributed exponentially. If a car has to stop it then acquires its own processing time.

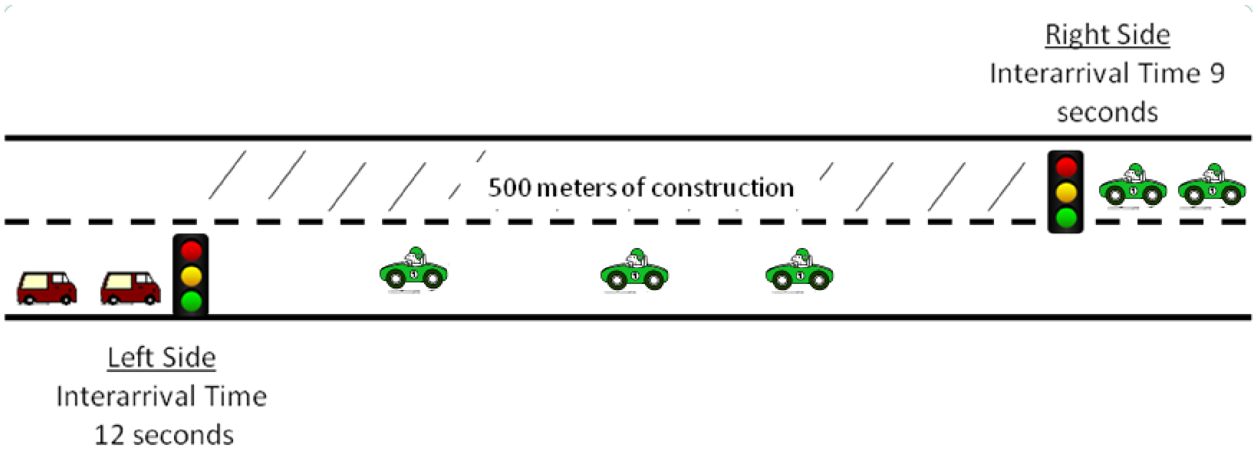


Figure 4: Graphical view of the simulated roadway.

A full cycle is where direction one, for descriptive purposes labeled the *right side* has a green light, then both sides are red to allow for the lane to clear, then direction two, labeled the *left side* has a green light, followed by another time period where both sides have a red light so the lane can clear. Again, when one lane has a green light to use the road, the opposite side has a red light.

Arnold, Nozari, and Pegden analyze the waiting time of all cars using the time that the light remains green for each side as their factors, ranging from 50 to 70 seconds each. Their control variates are the interarrival time of each side, the right side being 9 seconds and left side being 12 seconds. The processing time is held to a constant 2 seconds per car that requires it, and the response is to be the average waiting time of all cars that use the road over a one hour long time period. Arnold, Nozari, and Pegden received results showing that the left side is an effective control variate, while the right side is ineffective and the combination of both sides is less effective. They also use a central composite design requiring 13 runs to get this information. As stated in the earlier literature review, central composite designs are optimal designs in the D, A, and I components. With only two factors it is easy to get a design that is optimal in multiple criteria. A table of design values used for this original simulation are listed in Table 5.

Table 5: Values used by Arnold, Nozari, and Pegden.		
Variable	Value	
Response	Average waiting time of all cars through the system in 1 hour.	
Factor 1: Green Left	Low value of 50 Seconds	High value of 70 Seconds
Factor 2: Green Right	Low value of 50 Seconds	High value of 70 Seconds
CV 1: Mean Interarrival Time Left	12 seconds	
CV2: Mean Interarrival Time Right	9 seconds	
Time the Light is Red for Both Sides	55 seconds	

3.6.1.2 Research Adaptation

This research will reconstruct the scenario using the simulation environment ARENA and increase the number of factors and constraints so that the benefit of optimal designs can be investigated. This variation will show if any other factors should be included in the model while still investigating the impact of the same two control variates, interarrival time of the right side and interarrival time of the left side.

As shown in Table 1, *p* represents the number of factors and *q* the number of control variates. The expanded experimental design will use the original factors of the time that the light is green from the right and left, but will also use as additional factors the amount of time it requires to process a waiting car from the left and right directions, and the length of time both lights are red for the roadway to clear. Using a screening process we found the following five interactions of interest and will be used in the models as well: Green Light Time Right and Processing Time Right, (Green Light Time Right)^2, Green Light Time Right and Green Light Time Left, Green Light Time Left and Processing Time Right, and finally Processing Time Right and Red Light Time. These factors can be seen in Table 6. The control variates, *q*, will again be the interarrival time from the right and the left minus the mean interarrival time used in the respective exponential distribution. The total time for a cycle is used as a constraint on the factors by ensuring the sum of the green light time from both directions and the red light time is within a specific interval. The response remains that used by Arnold, Nozari, and Pegden, the average waiting time of all cars that cross the single lane section of road over a one hour time period.

3.6.1.3 Assumptions

The traffic model includes several assumptions relating the model to the real world scenario. This includes the fact the model starts empty. This reflects the assumption that it models traffic becoming busy, similar to a midday road just before rush hour. I am also assuming that there is no difference between vehicle types. Semi-trucks and small cars have the same acceleration after being stopped.

Table 6: The components of the simulation in this adaptation of the traffic simulation.

Variable	Value
Response	Average waiting time of all cars through the system in 1 hour. (Measured in seconds)
Factor 1: Green Left	[50, 70]
Factor 2: Green Right	[50, 70]
Factor 3: Processing Left	[1.5, 2.25]
Factor 4: Processing Right	[1.5, 2.25]
Factor 5: Red Light Time	[45, 55]
CV 1: Mean Interarrival Time Left	12 seconds
CV2: Mean Interarrival Time Right	9 seconds
Time the Light is Red for Both Sides	55 seconds
Constraint	-2 < Green Left + Green Right + Red Time < 2

3.6.1.4 Procedures

Using the factor levels shown in Table 6 for the overall design, JMP 10 was used to create optimal designs according to the D, I, and Alias criteria. The designs for each criterion and their corresponding efficiency measures, which are easily found using JMP, will be shown in the analysis section. These designs would be very difficult to compute by hand so the ease of use when computing them in JMP is a huge benefit. Due to the requirement for more runs than one plus the total number of factors and control variates, a 16 run design was chosen in order to allow for estimation of interactions beyond the main 5 factors. A modified half fractional design will be used for comparison methods as well. In this 2^{5-1} design the red light factor is aliased with the interaction GreenLeft*GreenRight*ProcessLeft*ProcessRight. In order for it to also follow the constraints that were used to create the variance optimal designs, the 4 points that violated the constraints were replaced with 4 center points.

Using Arena 13.9 a model depicting the scenario was created. The process analyzer tool allows the user to set up and run multiple replications with different parameters and then outputs the desired responses at the end of each replication. This tool is used to run all 16 design points easily while outputting the resulting response and control variate values for each individual run. The model can be seen in

Figure 5 showing each individual node that makes up the complete simulation.

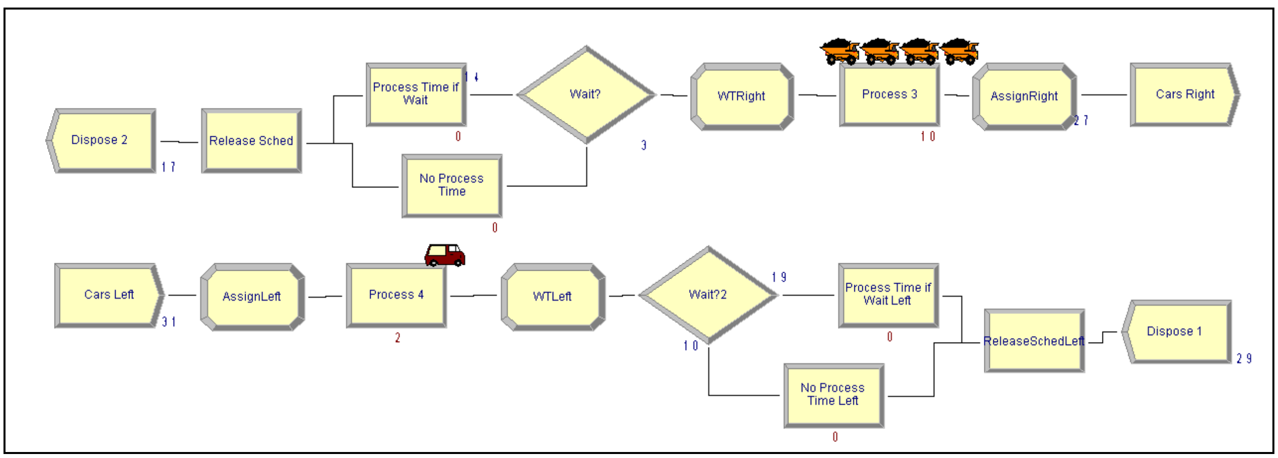


Figure 5: Screenshot of the ARENA model used for the traffic simulation.

3.6.2 618th TACC Simulation

3.6.2.1 Simulation Description

This simulation is a real world example to investigate manpower requirements from the 618th TACC flight management division (XOCM). Data was provided by a subject matter expert familiar with the unit and the process. Figure 6 is a flowchart for the flight planning process that the simulation is meant to capture. Each node is numbered and represents a process that someone at the unit must undertake. The unit is only interested in investigating the nodes which require a flight manager, the worker responsible for taking a sortie from the planning stage to take-off ready. Therefore the simulation is designed to capture nodes 7 to 23. The unit is comprised of 99 individuals, working 1 out of the 3 shifts, each lasting 9 hours. Shifts are meant to overlap to assist with changeover briefs which last 30 minutes. Because these people are committed to their responsibilities, they will not leave while they are currently working on a sortie. They will finish the process before departing for the day. However, they do need their breaks, which last 10 minutes and occur twice during their shift, as well as a 20 minute lunch break. While a flight manager's primary duty is to organize a sortie, one flight manager during each shift will work as the phone person, or ATM. This person will spend approximately half their day working the phone, and the other half planning sorties like any other flight manager. Due to health concerns, the unit does not wish to not overwork their flight managers if at all possible, but they must also accomplish their mission and get the flight planning done in time for the sortie to be flown. The subject matter expert gave ranges for values as well as the general idea of their interest in sensitivity analysis. From there they gave us control on creating the actual specifics for this study.

3.6.2.2 Responses

The unit is mostly concerned with two separate responses. In order to keep their flight managers happy and avoid over-working them, they want to investigate their utilization rate. This will be the rate at which they are performing flight planning duties with a sortie compared to the total time they are scheduled for potential work. They believe being constantly pushed to the edge would be undesirable and therefore would like to reduce the utilization rate of the flight managers.

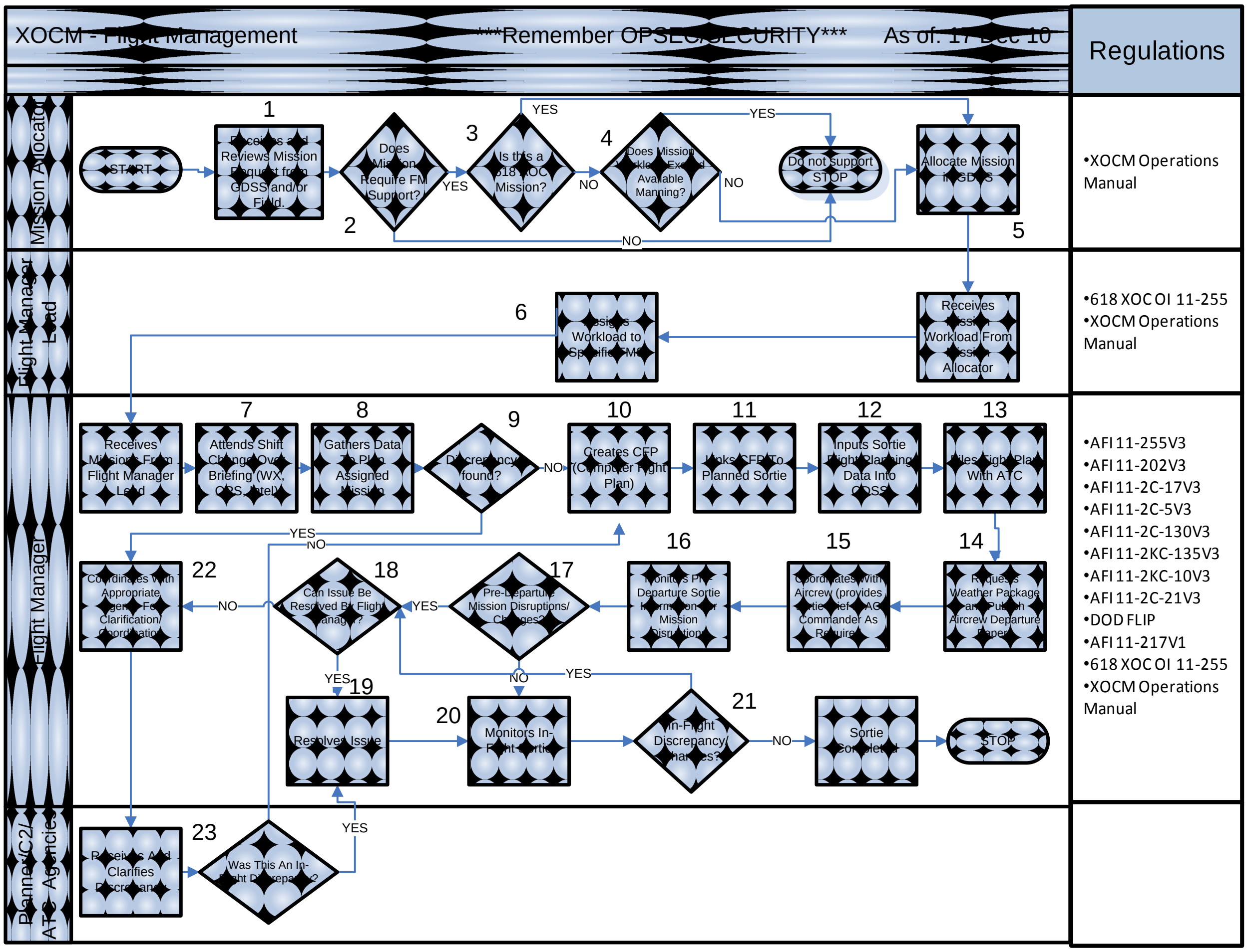


Figure 6: Flowchart for a single sortie.

The second response of concern is the time the sortie is actually in the planning process. In order to keep the flight squadrons happy, they want the planning process to take as little time as possible so that the TACC is not the reason a sortie cannot be completed on time. Although some delays are out of the TACC’s control, they believe the processes listed in the simulation could be of interest.

3.6.2.3 Variables

Several variables are considered by the TACC as potential factors that could be adjusted to achieve the desired responses. First, although the squadron size is defined, the number of flight managers working each shift can be adjusted. Therefore, the TACC is interested in the impact of scheduling between 21 and 23 flight managers per shift, not including the ATM flight manager.

The second variable of interest is the number of sorties scheduled per day. From the information provided by the subject matter expert, the mean number of sorties ranges from 210 to 240 and distributed normally with a standard deviation of 30 sorties.

The third variable is the probability that a discrepancy is found after gathering the data to plan the sortie. This is node 9 on the flowchart. This probability ranges between 25% and 35%.

The ATM flight manager is supposed to spend approximately half his day working the phone or other random tasks that take away from flight manager duties. Therefore, the TACC is concerned with the impact of adjusting this time to 40% to 60% of his day.

The final variable of interest is the rate at which an in-flight discrepancy is found, node 21. From information provided by the subject matter expert, we estimate that the probability of this occurring ranges from 35% to 40%. These factors are shown with the control variates and responses in Table 7.

3.6.2.4 Control Variates

There were several options for control variates in this system as many of the processes completed by the flight manager have an associated distribution provided by the subject matter expert and used in the simulation. However, there are 5 that are thought to be most correlated with the responses of interest. Because of the low cost of monitoring and applying control variates, all potential variates could have been used just to see their impact but that was determined to be unnecessary.

The first control variate is the average time it takes a flight manager to file flight plan with Air Traffic Control (ATC), node 13. The subject matter experts states that this process is exponentially distributed with a mean value of 15 minutes. Because taking longer at a single step means the process takes longer and the flight manager is busy longer, this process is thought to be connected to both responses.

The second control variate is the average time it takes a flight manager to resolve an issue when they have the ability to do so, node 19. This process has a mean value of 17 minutes, exponentially distributed. With the same reasoning as the first control variate, this should be correlated to both responses

The next control variate measures the time it takes a flight manager to coordinate with the appropriate agency for clarification or coordination when a flight manager cannot resolve a discrepancy, node 22. This is provided by the subject matter expert as another exponentially distributed value with a mean of 20 minutes. Although this only occurs when discrepancies are found and the flight manager cannot resolve them, the increased time and standard deviation it takes to complete could impact both responses significantly.

The fourth control variate is the time it takes a flight manager to create the computer flight plan, link it to the planned sortie, and input flight planning in GDSS. Although this is 3 separate steps, the subject matter would like it to be modeled as a single, exponentially distributed model with a mean value of 30 minutes. The increase time and variance within this process should impact both responses.

The final control variate measures the probability of pre-departure mission disruptions and changes, node 17. This is modeled as occurring 35% of the time, however due to the variation of the simulation, it may occur more or less frequently than this. This difference will be tracked and used as a control variate. Because this process means a sortie goes through several additional steps or not, it could certainly have a large impact on both responses. This control variate, as well as the previous 4, can be seen in Table 7.

3.6.2.5 Constraints

Two constraints were created for the simulation. The first constraint deals with the probability of a discrepancy being found initially and the probability of an in-flight change. We believe that if one probability approaches its maximum value, then the other discrepancy will not approach its maximum probability. The opposite is also true, if one probability nears its minimum value, the impact on the other probability will keep it from approaching its minimum value. For this reason, this constraint is modeled that the sum of the probability of an initial discrepancy and twice the probability of an in-flight discrepancy be between 97 and 113.

The second constraint involves the mean number of sorties per day and the adjustment made to the time the ATM flight manager spends away from flight manager duties. This constraint comes from the idea that on the days when sorties reach their maximum value, the ATM person will spend less time on the phone because they will be needed planning sorties. Also, on days when the number of sorties is low, the ATM flight manager will spend more time performing other duties. Because of the difference in units, this constraint is kept in coded space where each variable goes from -1 to 1. The sum of these two variables must then remain between -1.7 and 1.7. These constraints are shown in Table 8.

Table 7: Table of values used to create the simulation		
Variable	Value	Distribution
Response 1	Average Time in System for all Aircraft for a week in Minutes	
Response 2	Average Utilization rate for all Flight Managers for a week	
Factor 1: Number of FMs per Shift	[21, 23]	Constant
Factor 2: Mean Number of Sorties per Day	[210, 240]	Normal (σ =30)
Factor 3: Prob Discrepancy Found Initially	[25%, 35%]	Probability
Factor 4: ATM Phone Time Adjustment	[.8, 1.2]	Constant
Factor 5: Prob In-Flight Discrepancy/Change	[35%, 40%]	Probability
CV 1: Mean Time to File Flight Plan	15 minutes	Exponential
CV 2: Mean Time for FM to Resolve Issue	17 minutes	Exponential
CV 3: Mean Time to Coordinate with Appropriate Agency	20 minutes	Exponential
CV 4: Mean Time to Input Flight Plan Data	30 minutes	Exponential
CV 5: Prob of Pre-Departure Changes	35%	Probability
Changeover Brief	30 minutes	Constant
Mean Gather Data Time	20 minutes	Exponential
Mean Receive/Clarify Discrepancy w/ Agency Time	20 minutes	Exponential
Requests Weather Permits	5 minutes	Constant
Coordinates with Aircrew	1 – 2 minutes	Uniform
Prob FM can Resolve Issue	30%	Probability

Table 8: Constraints used to Create Optimal Designs in Coded Space
97 < Prob Discrepancy Found Initially+ 2*Prob In-Flight Disc < 113 (Natural Variables)
-1.7 < Mean Number of Sorties per Day + ATM Phone Time Adjustment < 1.7 (Coded Variables)

3.6.2.6 Assumptions

The TACC simulation includes some assumptions that are similar to the traffic simulation in terms of dealing with the methods used to create and evaluate the system. However, the simulation also includes assumptions which were used to take the real world scenario to the simulation world. This includes the assumption that the changeover brief occurs for exactly 30 minutes at the beginning of each shift for all workers and the required breaks can be taken at once at the beginning of the shift rather than spread out throughout the day without impacting results because they are taken when the worker is not currently assigned an aircraft. We assume that all times and distributions provided by the 618th TACC accurately represent the true system. The model also has the assumption that workers use a 24 hour rotation for a 5 day work week with no consideration for vacation time, work time, or other duties which are to be performed by the worker such as computer based training, fitness tests, etc. As previously discussed, the simulation forces flight managers to finish the aircraft they are working on before leaving for the day. Finally, because the ATM flight manager splits duties between managing the phone and typical flight manager duties, although the ratio may vary, the simulation assumes that he spends exactly the amount time determined by the design for that specific run.

3.6.2.7 Procedures

The procedures for this simulation are very similar to those described in the earlier traffic simulation. Now that the response, variables, control variates, and constraints have been determined in Table 7, JMP can be used to create the 3 optimal experimental designs: D, I, and Alias. These designs will be used alongside a full factorial model which ignores the constraints. The designs will be run through the TACC ARENA simulation, shown in Figure 7, using the ARENA Process Analyzer. In addition to the processes used for the traffic simulation we examine the correlation between the response and control variate values. In addition, for the TACC simulation, a third model is created with control variates entered directly into the original model developed without control variates. The process of creating models will be conducted for all 4 designs and both responses, creating a total of 24 models. While each design will have a separate correlation matrix, 4 in total, each of the 24 models will have its own measures of efficacy of R^2adjusted, significant coefficients, coefficient half widths, and variance estimates.

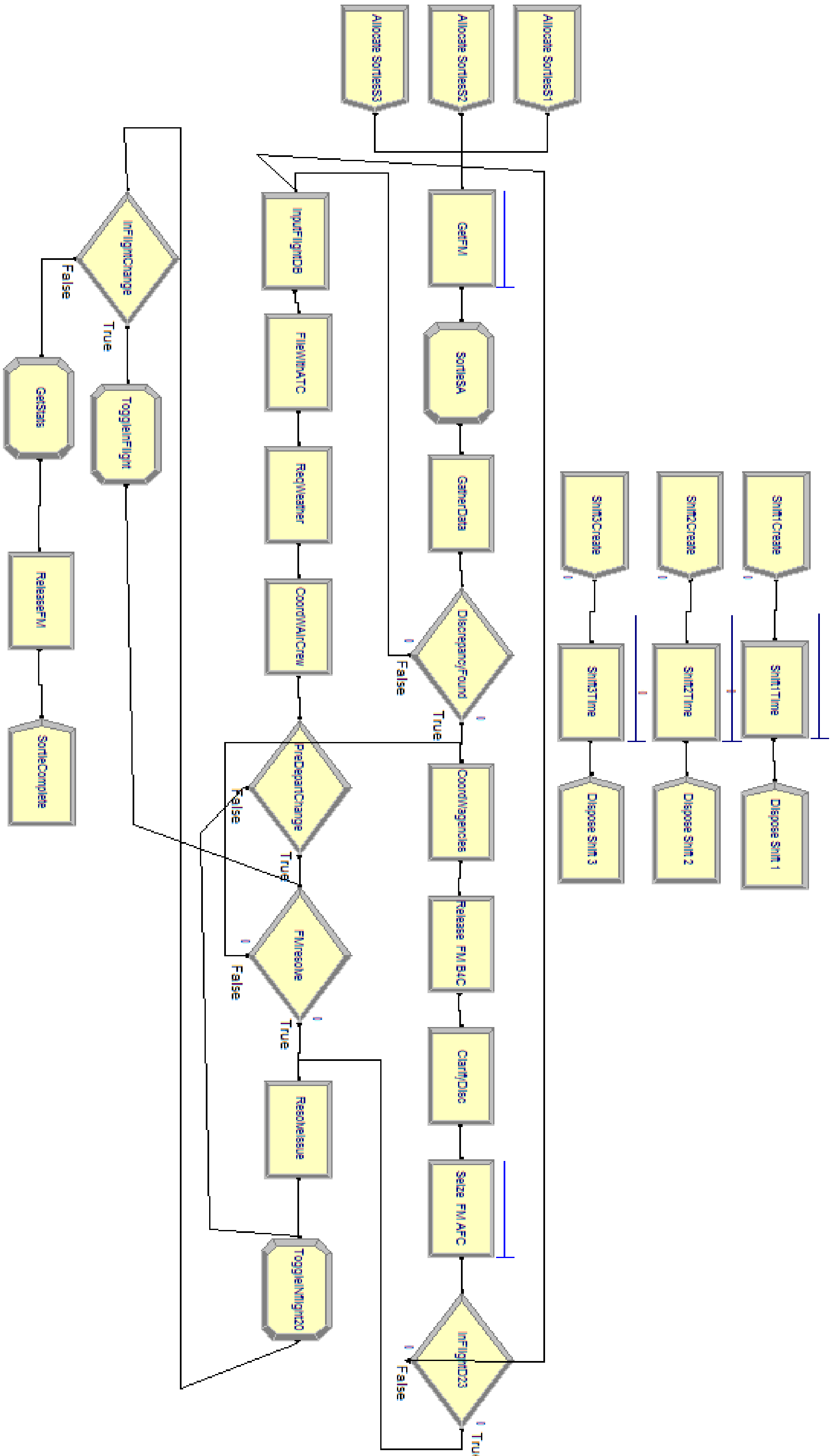


Figure 7: Screenshot of the ARENA model used for the TACC simulation.

For this simulation, calculating the prediction measures of effectiveness for each model will be performed with 25 random points. Predicted mean square error, prediction half width, and coverage will be used to judge the efficacy of these control variates. Again, comparing the performance of the models in the end, to what we knew about the model at the beginning allows us to develop recommendations on when control variates should be employed in real world scenarios and when their benefit may not be seen.

3.7 Summary

Two simulations will be used to demonstrate the benefits of control variates. Four experimental designs will be created, a variation of a half fraction, D-Optimal, I-Optimal, and Alias optimal. The simulation output for each design will be recorded and processed to create statistical measures to evaluate the potential benefits of both the experimental design and the variance reduction from using control variates. Each method has assumptions associated with it which if not true could impact the expected theoretical benefits gained by using these methods. This is clearly not an exhaustive use of the techniques as limitations exist. But again, it is a great start towards investigating the impact that control variates have on optimal designs and their use with evaluating simulations.

4. Analysis and Results

4.1 Chapter Overview

After the factors, control variates, and responses have been decided, the optimal designs have been found, the simulation has been created, and the simulation run to get results; these results must be analyzed to determine the effect of the techniques used. This section details what was found through this research and how it can be interpreted and used in the future. The traffic lane simulation and the Air Force flight management unit simulation are very different scenarios that both offer insight into the use of control variates and optimal designs for metamodeling.

4.2 Traffic Lane Simulation

The factors, control variates, constraints and response of interest listed in Table 6 were input into JMP to calculate a 16 run optimal design for each of the following criterion: D-optimal, Alias optimal, and I-optimal. Combined with the half-fraction design described in the methodology, the four designs of interest shown in Table 9, Table 10,Table 11, and Table 12 were created. Again this design used the initial design of all 4 factors and 4 interactions. As the results will show, this first model shows how control variates are not always effective. Because of the long queue length these control variates are ineffective, additional iterations of the model will show that the benefits of control variates are based on each individual case.

Table 9: Half Fraction Design

GreenLeft	GreenRIght	ProcessL	ProcessR	RedTime
-1	-1	-1	-1	1
1	-1	-1	-1	-1
-1	1	-1	-1	-1
0	0	0	0	0
0	0	0	0	0
1	-1	1	-1	1
-1	1	1	-1	1
1	1	1	-1	-1
0	0	0	0	0
1	-1	-1	1	1
-1	1	-1	1	1
1	1	-1	1	-1
-1	-1	1	1	1
1	-1	1	1	-1
-1	1	1	1	-1
0	0	0	0	0

Table 10: Design and efficiency values for D-Optimal Design

D Efficiency	93.6321
G Efficiency	96.7589
A Efficiency	90.989
Average Variance of Prediction	0.17463
Design Creation Time (seconds)	5.07

GreenLeft	GreenRIght	ProcessL	ProcessR	RedTime
-1	-1	1	-1	1
1	-1	-1	-1	1
-1	0	-1	-1	-1
-1	-1	-1	1	0
1	1	-1	1	-1
-1	1	1	-1	1
1	1	1	1	-1
1	-1	1	1	1
-1	1	1	-1	-1
1	1	-1	-1	-1
-1	1	-1	1	1
-1	1	1	1	1
1	-1	1	-1	-1
-1	-1	-1	1	0
1	-1	1	1	-1
1	0	-1	-1	1

Table 11: Design and efficiency values for I-Optimal Design	
D Efficiency	92.894
G Efficiency	96.24581
A Efficiency	91.94224
Average Variance of Prediction	0.173285
Design Creation Time (seconds)	5.85

GreenLeft	GreenRIght	ProcessL	ProcessR	RedTime
1	-1	-1	-1	-1
1	0	-1	-1	1
1	-1	1	1	-1
-1	1	-1	-1	-1
-1	1	-1	1	1
-1	1	1	1	1
-1	-1	-1	1	0
-1	1	1	-1	-1
1	1	-1	1	-1
-1	-1	1	-1	1
1	1	1	1	-1
1	0	-1	-1	1
0	1	1	-1	1
0	-1	1	-1	-1
-1	-1	-1	1	0
1	-1	1	1	1

Table 12: Design and efficiency values for Alias Optimal Design	
Alias Optimal Design	
D Efficiency	90.76591
G Efficiency	92.02592
A Efficiency	89.49007
Average Variance of Prediction	0.1768
Design Creation Time (seconds)	591.0167

GreenLeft	GreenRIght	ProcessL	ProcessR	RedTime
-1	-1	-1	1	0.1
-1	-1	1	-1	1
-1	-0.74	-1	1	-0.26
-1	1	-1	-1	-1
-1	1	-1	-1	1
-1	1	1	1	-1
-1	1	1	1	1
-0.609	-0.432	1	-1	-0.959
0.608	0.392	-1	1	1
1	-1	-1	-1	-1
1	-1	-1	-1	1
1	-1	1	1	-1
1	-1	1	1	1
1	0.729	1	-1	0.271
1	1	-1	1	-1
1	1	1	-1	-0.1

Each row in these experimental design tables represents a different design input into the ARENA model with the columns containing the factor values unique for the design point. ARENA, with the help of Process Analyzer, provide as output the average waiting time, average interarrival time for the right side, and average interarrival time for the left side for each design point in the design. The output for the experimental designs can be found in Appendix B.

The process analyzer output and the design structures from JMP were then used to compute the measures of effectiveness with the MATLAB code in Appendix A. The variance estimate, coefficient values, and 90% confidence half width for each coefficient calculated by the MATLAB code are found in Table 13 and Table 14.

The initial results of the traffic lane simulation were not as expected. I anticipated a clear benefit from at least one control variate in each design shown by Arnold, Nozari, and Pegden (1984). However, none of my four designs showed any improvement in variance estimation and half width due to control variates. The results for variance can be found in Table 13 with half-width results found in Table 14. Column 2 of Table 13 is the variance estimate for each model. Column 3 is the variance estimate multiplied by the loss factor to calculate the adjusted variance of the coefficients. Column 4 shows the adjusted variance multiplied by the associated *F* statistic to calculate the squared half width of the coefficient confidence interval. These are shown because a small reduction in variance estimation may be insufficient to overcome the loss of degrees of freedom due to the need to estimate the covariance of the response and the control variates.

As can be clearly seen in the Table 13, only one model showed a reduction in variance estimation once degrees of freedom were accounted for. The single exception is when control variate one was applied to the full factorial model. Although the initial variance estimate has

Table 13: Estimates of variance and measure of efficacy upon employing different control variates.			
		$\hat{\tau}^2$	$\hat{\tau}^2 * \frac{n-p-1}{n-p-q-1}$
Half Fractional			
	No CV	653.9235	653.9235
	CV 1	489.3009	611.6261
	CV 2	637.4541	796.8176
	CV 1 and 2	495.6387	826.0644
D - Optimal			
	No CV	207.4824	207.4824
	CV 1	230.6051	288.2564
	CV 2	232.627	290.7838
	CV 1 and 2	282.5526	470.921
Alias Optimal			
	No CV	337.7144	337.7144
	CV 1	301.6191	377.0239
	CV 2	385.6535	482.0669
	CV 1 and 2	376.8746	628.1244
I - Optimal			
	No CV	47.01747	47.01747
	CV 1	43.46899	54.33623
	CV 2	48.97203	61.21504
	CV 1 and 2	53.86331	89.77218
Column 1	Column 2	Column 3	Column 4

been reduced in some cases, when the new variance estimate is used to estimate the variance or half width of the coefficient, the value would be larger using control variates than without them as seen in columns 3 and 4. The closest control variate to showing benefit was control variate one across all designs, the interarrival time of cars from the right. This trend can be seen in all optimal designs although not enough to make its inclusion in the model worthwhile. The reduction in variance was not enough to counteract the lost degrees of freedom for any of the optimal designs checked. As stated earlier not all control variates work effectively. However, the cost of adding these control variates and getting the data after the simulation run completed was so minimal that the potential for benefit as seen in the one design is still worthwhile. This also shows that control variates affected each design the same. Although it did not show benefit in any of these designs, the trend that control variate one outperformed control variate two was universal and gives an analyst information that the interarrival time from the right has more correlation to the waiting time than does the interarrival time from the left. All of the optimal designs outperformed the half-fraction design when using variance as the measure of effectiveness with the I-optimal design achieving the lowest variance estimate. Clearly estimated variance also plays a part in the half-width on each model coefficient. These can be compared in Table 14 below, but only the no control variate model and the model including control variate one are used because control variate one already proved to be the best control variate option. As well as having the smallest variance estimate, the I-optimal design also has the smallest half-width on the coefficients, followed by the half fraction and the D-optimal design.

Table 14: 90% simultaneous half widths for each coefficient

	Half-Width							
	Half Fraction		D – Optimal		Alias Optimal		I - Optimal	
	No CV	CV 1	No CV	CV 1	No CV	CV 1	No CV	CV 1
Intercept	51.555	49.860	44.075	54.531	63.978	70.956	33.829	38.173
GL	36.455	35.256	17.224	21.310	19.56	21.696	11.941	13.475
GR	36.455	35.256	16.555	20.482	20.547	22.788	8.5505	9.6485
PR	36.455	35.256	16.000	19.795	18.584	20.611	8.5505	9.6485
Red	29.765	28.786	20.424	25.269	21.425	23.762	9.6258	10.861
GR*PR	36.455	35.256	16.555	20.482	20.812	23.082	8.5505	9.6485
GR*GR	36.455	35.256	47.843	59.193	75.018	83.200	37.820	42.676
GL*GR	60.453	58.466	19.011	23.521	23.318	25.862	13.39	15.119
GL*PR	31.571	30.533	17.224	21.310	19.868	22.035	10.523	11.875
PR*Red	36.455	35.256	20.424	25.269	24.193	26.832	14.031	15.833

The prediction power of each model is also used to measure the effects of control variates. Because the addition of control variates makes adjustments to the coefficients of the other factors, the use of control variates could potentially offer better prediction without a better variance. Only the non-control variate model and the model using control variate one were used for the prediction methods. The interarrival time for the right side clearly outperformed the other control variate combinations making it the best one to explore. The randomized points and their results for prediction error can be found in Table 15.

Table 15: Random prediction points and the prediction mean square error of each model.

	GreenLeft	GreenRIght	ProcessL	ProcessR	RedTime	True Response
1)	0.255	-0.755	0.061	-0.927	0.629	73.215
2)	-0.725	0.602	-0.144	0.679	-0.733	68.389
3)	-0.083	0.035	-0.755	-0.839	0.694	72.72
4)	-0.209	-0.929	0.067	0.244	-0.815	72.054
5)	0.074	0.034	0.893	-0.224	-0.294	70.77
6)	-0.499	-0.972	-0.693	0.599	-0.977	76.471
7)	-0.619	0.347	0.445	0.613	-0.692	69.466
8)	0.336	-0.388	-0.725	0.619	0.579	90.896

NO CV	MSE	CV1	MSE
Half Fraction	331.451	Half Fraction	371.919
D-Optimal	464.568	D-Optimal	424.860
Alias-Optimal	123.309	Alias-Optimal	140.084
I-Optimal	870.926	I-Optimal	969.459

In this scenario the only control variate model to outperform its counterpart is the D-optimal design. However, the alias optimal design presented the best prediction values as it shows the smallest mean square error. The I-Optimal design performs by far the worst of those tested even though it had the smallest variance estimate. This case shows only a single replication so some of the results could be due to bad luck, and randomly getting results that perform poorly. However, because these techniques are used when money and resources are limited, the number of runs and replications are also limited, this single run method is most like the actual application of the techniques.

4.2.1 Additional Models

Because the results were not as expected, I chose to investigate why and possibly modify the simulation and perform the analysis again to see what may come from slightly different parameters. Upon further inspection, the factors chosen created quite a large queue of cars waiting to use the roadway. We should have been more cautious of queue length sensitivities after these results. The effect of interarrival time would be considerably diminished when cars begin to wait so long that the difference in interarrival time from the mean is negated by the long wait. Although the interarrival time has some impact on causing the large queue, it is also the factors of the time the light is green, the processing time, and the red light time causing the queue to grow. When the factors create a system that cannot process as many cars in a cycle as the number of cars arriving into the cycle, there is no choice but for the queue to grow infinitely large.

4.2.2 Model Modification 1

Table 16: Values used for Modification 1.

Variable	Value
Response	Average waiting time of all cars through the system in 1 hour.
Factor 1: Green Left	[50, 70]
Factor 2: Green Right	[50, 70]

Replication 1		Replication 2	
CV 1: Mean Interarrival Time Left	12 seconds	CV 1: Mean Processing Time Left	2 seconds
CV2: Mean Interarrival Time Right	9 seconds	CV2: Mean Processing Time Right	2 seconds
Time the Light is Red for Both Sides	25 seconds	Time the Light is Red for Both Sides	60 seconds

Therefore, to investigate the relationship between queue length and the control variates, the experiment was run again. The benefits of control variates can be specific to the design space being investigated. Although a single simulation scenario may show no benefits, another application of control variates on the same simulation may be more beneficial. However, this time the red light time was held to a constant for each design, but changed from one design to another to compare the difference the shorter queue may have on the effect of control variates. Also, to investigate the impact of using interarrival time versus service time as the control variates, we reverted back to the original design presented by Arnold, Nozari, and Pegden. In their example only green light time from each side was used as a factor in a 13 run central composite design. The values for these replications of the design can be found in Table 16. Table 17 shows the results of reducing the red light time to 25 seconds while using interarrival times as control variates, and then increasing red light time to 60 seconds while using processing time of each side as the two control variates.

Table 17: Comparison of variance estimates when changing the control variate and the factor Red Light Time.

Red Light = 25 Sec Control Variate = Interarrival Time			Red Light = 60 Sec Control Variate = Processing Time		
	$\hat{\tau}^2$	$\hat{\tau}^2$(Loss Factor)		$\hat{\tau}^2$	$\hat{\tau}^2$(Loss Factor)
No CV	9.743857	9.743857	No CV	357.0402	357.0402
CV 1	1.573996	1.888795	CV 1	172.8528	207.4233
CV 2	7.639409	9.167291	CV 2	406.4422	487.7306
CV 1, 2	1.093835	1.640752	CV 1, 2	202.8868	304.3303

This table clearly shows that the change in red light time affects the benefit of each control variate. A shorter time for the red light means the cycle is shorter and cars do not wait as long, limiting the chance of a long queue building up and making the interarrival time more correlated with the response and offers more benefit in reducing the variance. Inversely, the longer red light time forces cars to wait longer, increasing the chance of longer queues and making the benefit of using processing time as the control variate more clear.

This result highlights the importance of the earlier assumption of constant variance throughout the design. In our earlier designs red light time was a factor, and according to the assumptions for using control variates, there is a constant variance for all design points. However, this latest finding makes that a poor assumption because the covariance between the response and the control variate changed as the factors changed. This shows the importance of ensuring the assumptions are supported and true within every model and are considered when choosing factors for the design. Because the length of time that the light was red had an interaction effect on the control variate, the effectiveness of control variates were limited by our design space.

4.2.3 Model Modification 2

Table 18: Values used for Modification 2.

Variable	Value
Response	Average waiting time of all cars through the system in 1 hour. (Seconds)
Factor 1: Green Left	[45, 70]
Factor 2: Green Right	[55, 80]
Factor 3: Interarrival Left	[11, 13]
Factor 4: Interarrival Right	[9, 11]
CV 1: Mean Processing Time Left	2 seconds
CV2: Mean Processing Time Right	2 seconds
Time the Light is Red for Both Sides	60 seconds
Constraint 1 – Coded Space	-1.8 < Green Right + Green Left < 1.8
Constraint 2 – Coded Space	-1.8 < Interarrival Right + Interarrival Left < 1.8

Now that I have the knowledge that when red light time is high, 60 seconds, the processing time should be used as the control variate the original experiment will be redone. Again, exploring a different part of the space could results in different benefits from control variates. The first modification showed no benefits. But this modification uses the same simulation, just adjusting the factors of interest, and benefits may be more explicit. This highlights the need for the control variates to correlate with the response, which can change from one scenario to the next. This variation will have the red light time set to a constant 60 seconds, the processing time from the right and left will be used as the two control variates, and the green light of each side as well as the interarrival time will be used as the four factors. Again, four designs will be constructed: full factorial (does not require constraints), D-optimal, Alias optimal, and I-optimal. Now that there are only four factors, the optimal designs will be constructed with twelve runs each and the new constraint that Green Time Right + Green Time Left rests within the coded interval of (-1.8, 1.8), as does the Interarrival Time of the Right side + the Interarrival Time of the Left side. These constraints eliminate part of the design space, forcing the optimal criteria design software to find the best way to allocate the design points to meet the design criteria in the constrained space. The values used for this modification are shown in Table 18. In this experiment, the prediction mean square error of each design will also increase from predicting eight points to predicting twenty. Another difference in this modification is that no predetermined model will be used. Rather, all eight models, four designs with a no control variate and a control variate model each, will be created using the JMP stepwise function. This will be done to recreate what an analyst would conclude if he were attempting to analyze the data when control variates were not available, and compare it to what he would conclude if the control variates were available.

The results of this modification will be analyzed similar to the initial model. The variance estimates of each model will be compared. Then the factors and half-widths of each model can be compared to see what factors would be found insignificant because the variation explained by control variates is masking their importance. Then the prediction ability on the twenty points is investigated to see if the model found using control variates can better predict random points in the design space.

Table 19: Required variance reduction for future benefits in average waiting time of Modification 2				
Design	Model	# of Factors	# of Control Variates	$\frac{n-p-q-1}{n-p-1}$
Full Factorial	No CV	4	0	
	CV Always Available	9	1 (CV2)	0.83
D – Optimal	No CV	5	0	
	CV Always Available	6	2 (CV 1 and 2)	0.60
Alias Optimal	No CV	5	0	
	CV Always Available	7	1 (CV1)	0.75
I – Optimal	No CV	5	0	
	CV Always Available	8	1 (CV1)	0.67

4.2.3.1 Variance Results

The simulation results were input into JMP and 8 models were created, 2 for each design structure. The number of factors and significant control variates, along with the largest value that $\frac{\tau^2}{\sigma^2}$ can be to see benefits of control variates, are shown in Table 19. The variance estimates for the model found using control variates is shown in Table 20. The $\hat{\tau}^2$ is the variance estimate while the third and fourth columns are adjusted for the number of control variates used to compute the variance and half width estimates. The decreased variance must not be overshadowed by the lost degrees of freedom from the additional number of variables. The results in Table 20 clearly show the benefits achievable using control variates. While the full factorial model shows the left side processing time to be the most beneficial, the D-Optimal design shows the use of both processing times to be the most beneficial use of control variates, and the I-Optimal models show processing time from the right to be the most beneficial control variate for variance estimation. The Alias Optimal model is constructed in JMP where the processing time from the right is the most beneficial control variate, while the variance estimate shows the use of both control variates to be slightly better. This proves the importance of including all possible control variates. While one may seem important in this run, the randomness that comes with variation may prove it to be non-beneficial the next replication while another control variate shows important.

Table 20: Variance estimates of the modified model upon employing different control variates

	$\hat{\tau}^2$	$\hat{\tau}^2 * \frac{n-p-1}{n-p-q-1}$
Full Factorial		
No CV	815.6238	815.6238
CV 1	915.9673	1068.628
CV 2	157.2807	183.4942
CV 1 and 2	142.885	200.039
D - Optimal		
No CV	366.6149	366.6149
CV 1	223.0362	278.7952
CV 2	249.4987	311.8733
CV 1 and 2	117.9199	196.5332
Alias Optimal		
No CV	711.5371	711.5371
CV 1	47.61876	71.42814
CV 2	888.9834	1333.475
CV 1 and 2	23.5704	70.7112
I - Optimal		
No CV	922.3359	922.3359
CV 1	11.52215	17.28322
CV 2	567.7847	851.6771
CV 1 and 2	8.022306	24.06692

4.2.3.2 Factor Coefficient and Half-Width Results

Due to the fact that each design creates two models independently, they often showed different factors to be significant. This can be attributed to the fact that the occurrence of variation in processing time masks the significance of truly significant factors. When the control variates are used to account for some of the variance in the system, the true significance of each factor can be tested. Table 21 clearly shows the difference in factor inclusion across the different models, where the coefficient listed shows also the presence of that factor. This table also shows the wide range that coefficient values can take even when they are included in different models.

Table 22, shown below, displays the half-width for each coefficient significant to the model using the equation $t_{\frac{\alpha}{2*p'},\ n-p-q-1}\sqrt{\frac{n-p-1}{n-p-q-1}\hat{\tau}^2(X'X)^{-1}}_{jj}$. As can be seen in direct comparison, each control variate model has a smaller half-width on the common significant factors. The I-optimal design with control variates has a half-width better than the full factorial model even though the full factorial model used four more design points.

Table 21: Coefficients for the terms included in each model								
	Coefficients							
	Full Factorial		D – Optimal		Alias Optimal		I - Optimal	
	No CV	CV	No CV	CV	No CV	CV	No CV	CV
Intercept	102.092	102.293	-171.44	103.365	107.011	103.810	106.724	104.479
GL		-5.291	0.269	5.767	-6.861	-3.166	14.550	19.636
GR	-8.471	-4.512		1.553	-2.206	6.326	-13.721	-27.566
Arr L	-11.625	-5.969	-6.431		-7.015	-2.468		-8.654
Arr R	-6.604	-7.853		-7.332		-17.270	-19.548	-16.721
GL * GL			381.651					
GL * GR		-7.533		-9.191				
GL * Arr L		3.559			13.645	20.352		
GL * Arr R		-3.408					-7.977	-9.457
GR * Arr L								7.476
GR * Arr R		3.667		9.175		10.070		9.105
Arr L * Arr L			-113.09					

4.2.3.3 Prediction Results

The prediction ability of a model is also of particular concern. As seen in the initial models, prediction was not improved by the use of control variates when they did not have a strong influence on the response. In this modification, the processing time for each side of cars has proven to have a strong relationship with the average waiting time of all cars in variance estimation and simultaneous coefficient half widths. To evaluate the prediction accuracy of the models, they are used to predict twenty random points in the design space. These twenty points

Table 22: Half width for the terms included in each model								
	Simultaneous Half Width							
	Full Factorial		D – Optimal		Alias Optimal		I - Optimal	
	No CV	CV	No CV	CV	No CV	CV	No CV	CV
Intercept	9.606	7.285	415.178	14.395	23.748	21.685	26.847	6.514
GL		7.285	14.385	15.108	23.907	21.794	28.592	7.154
GR	9.606	7.285		15.753	27.522	25.903	29.748	7.323
Arr L	9.606	7.285	17.446		25.315	22.841		7.587
Arr R	9.606	7.285		15.892		23.721	28.642	6.992
GL * GL			487.046					
GL * GR		7.285		16.539				
GL *Arr L		7.285			26.404	24.910		
GL * Arr R		7.285					31.300	7.604
GR * Arr L								8.726
GR * Arr R		7.285		17.368		26.554		7.981
Arr L * Arr L			115.087					

also satisfy the constraints used to create the optimal experimental designs. For these results, the process was used to create 4 prediction models of each design. The randomized points and the mean square error for each replication can be found in Table 23. In only 4 of the 16 replications did the model without control variates predict the points better than the model found using control variates. All four composite models, the models found using data from all four replications, showed a better prediction model using control variates for all design criteria.

4.2.3.4 Model Adequacy

There were many assumptions in using control variates and linear regression. A few of these discussed will now be checked and results compared. This will show whether these assumptions affect the benefit of control variates. These assumptions include constant covariance of the response and the control variates throughout the design, Constant residual variance across the prediction points, model specification, and the assumption of normality among the residuals.

The first assumption of constant variance and covariance can be easily shown to not be true in this design by comparing the covariance matrix of 250 replications at several design

Table 23: List of randomized prediction points and prediction results.

	GreenLeft	GreenRIght	InterArrival L	InterArrival R	True Response
1)	45.696	76.902	11.849	10.737	101.974
2)	46.014	66.972	12.872	10.250	87.133
3)	46.401	67.132	12.112	9.629	91.772
4)	49.744	71.635	12.977	10.438	85.118
5)	50.537	70.985	12.659	9.103	88.274
6)	51.803	73.166	12.801	10.531	85.256
7)	53.712	72.993	11.238	9.962	92.305
8)	54.738	74.966	12.535	9.319	87.484
9)	55.104	56.995	12.643	10.780	85.8935
10)	56.898	69.254	11.973	10.655	85.7335
11)	58.371	77.741	12.980	9.409	86.5245
12)	59.298	64.756	11.036	9.587	91.967
13)	62.153	75.529	11.103	9.595	90.242
14)	62.198	67.825	11.077	10.971	87.3845
15)	63.122	72.502	11.592	10.308	87.524
16)	65.762	66.807	12.448	10.736	86.273
17)	67.806	77.639	11.001	10.172	89.344
18)	67.889	56.061	11.781	10.025	97.9325
19)	69.212	64.174	11.721	10.988	86.1845
20)	69.539	75.524	11.867	10.028	88.4255

Design	MSE NO CV	MSE w/ CV	Difference	% Change
Full Factorial	59.40	54.98	-4.41	-7.42%
	53.77	89.95	36.18	67.29%
	286.19	110.33	-175.87	-61.45%
	62.78	55.40	-7.38	-11.76%
Full Factorial - Comp	70.74	56.98	-13.76	-19.45%
D-Optimal	45.64	51.93	6.29	13.78%
	71.47	52.90	-18.57	-25.98%
	71.85	53.50	-18.35	-25.54%
	116.48	71.86	-44.62	-38.31%
D-Optimal - Comp	41.53	32.67	-8.86	-21.33%
Alias Optimal	111.76	78.47	-33.29	-29.79%
	63.22	74.11	10.89	17.23%
	87.69	75.54	-12.15	-13.86%
	73.59	45.88	-27.71	-37.65%
Alias Optimal - Comp	136.42	40.31	-96.12	-70.46%
I-Optimal	82.73	28.83	-53.90	-65.15%
	102.97	94.29	-8.69	-8.44%
	151.75	63.99	-87.77	-57.84%
	79.67	128.61	48.95	61.44%
I-Optimal - Comp	78.14	22.63	-55.51	-71.04%

points. Table 24 shows the covariance matrix for 5 points throughout the design. It can be easily seen that the covariance between the response and the control variate 1 ranges from -.6372 to 3.1

and between -.55 and 3.116 for control variate 2. The covariance between the control variates and the variance of each control variate is fairly consistent. Meanwhile, the variance of the response ranges from 26.0953 to 7051.6076. These values clearly violate the assumption.

An assumption when using linear regression is that the residuals from the model are constant across the range of prediction values. When investigating this assumption it appeared to have a funnel, which would violate this assumption. Therefore, a Box-Cox transformation was done on the model which adjusted the response of Y to Y^{-2}. This transformation also covers the fact that this is a first order model when a second order model may be more appropriate.

Table 24: Covariance matrix for several points throughout the design

GreenLeft	GreenRight	InterArr L	InterArr R
-1	-1	1	1
	WT	CV1	CV2
WT	29.6943		
CV1	0.1950	0.0125	
CV2	0.1775	0.0001	0.0120

GreenLeft	GreenRight	InterArr L	InterArr R
-1	1	-1	1
	WT	CV1	CV2
WT	6483.0363		
CV1	-0.6372	0.0119	
CV2	3.1161	-0.0012	0.0137

GreenLeft	GreenRight	InterArr L	InterArr R
1	-1	-1	-1
	WT	CV1	CV2
WT	7051.6076		
CV1	3.1006	0.0095	
CV2	-0.5500	-0.0001	0.0142

GreenLeft	GreenRight	InterArr L	InterArr R
1	1	1	-1
	WT	CV1	CV2
WT	26.0953		
CV1	0.1512	0.0112	
CV2	0.1794	-0.0007	0.0141

GreenLeft	GreenRight	InterArr L	InterArr R
0	0	0	0
	WT	CV1	CV2
WT	29.6943		
CV1	0.1950	0.0125	
CV2	0.1775	0.0001	0.0120

However, after accomplishing the transformation and running the analysis again, the same results were found for variance reduction and prediction accuracy. Although a transformation was appropriate for the model, control variates displayed the same benefits with and without a transformation.

Another assumption mentioned earlier in this thesis is that the model is specified appropriately. While the transformation could make up for some of the need for a 2nd order model, a face-centered central composite design was checked to ensure that this model would not change the benefits of control variates. After again running analysis using the model design, factors, control variates, and response, the same benefits were seen with this design. The 2nd order terms showed to be significant when creating the best model. But applying control variates could improve this model in both variance reduction and prediction even further.

A fourth assumption to check is the presence of outlier points. Because this scenario is based on the waiting time of a queue, there is potential for the queue to “blow up”. This is when more people arrive than can be physically processed by the system, causing the waiting time for the replication to get extremely large compared to the other replications. This occurrence would create a non-linear design space, making it very difficult to model with a linear regression model. This could also be the reason for the nonconstant covariance mentioned in the first assumption. In order to check the effect of these outliers new constraints were put on the design points where it seemed like the system was blowing up most often. This included the points (1, -1, 1, -1) and (-1, 1, -1, 1) for the GreenLeft, GreenRight, ArrivalLeft, and ArrivalRight; respectively. There was also an adjustment that removed responses greater than 150 seconds as they appeared to be much greater than the mean of the other replications. These points appear to be problematic because they cause the queues to be unstable. This instability means the lights cannot process all of the cars that arrive and wait times grow extremely large. The graph of responses then began to approach a more linear model. This constrained model was then replicated 250 times at each design point.

While the previous optimal design and full factorial model showed no prediction benefit with control variates at 250 replications, these models included outlier points and design points causing nonconstant covariance. This constrained model removed these outlier design points and was analyzed similar to the previous models. The resulting model showed the benefit of control variates at only 1 replication all the way to 250 replications. As the number of replications increase the model should perform better than low replication models as there are more points to help create the model and therefore less variance remaining for the control variates to account for.

The figures below show the benefit as the number of replications are increased from 1 to 250. Figure 8 shows the benefit in variance reduction. The x-axis is the number of replications included in the model, while the y-axis shows the ratio of variance with no control variates divided by the variance with the control variate of interest. Therefore, the greater the value on the y-axis, the more benefit there is from using that control variate. Clearly there is a benefit of using both control variates or just control variate 1 throughout the entire space, while control variate 2 may be nonbeneficial at the start, as replications are increased, the control variate become even more beneficial than control variate 1.

Figure 9 shows the number of replications in the model on the x-axis while the y-axis shows the predicted mean square error. As replications increase they approach a similar result in mean square error. However, with less than about 150 replications it can be easily seen that using both control variates returns the best prediction model, and using only one of either control variate creates a better prediction model than the model with no control variates accounted for.

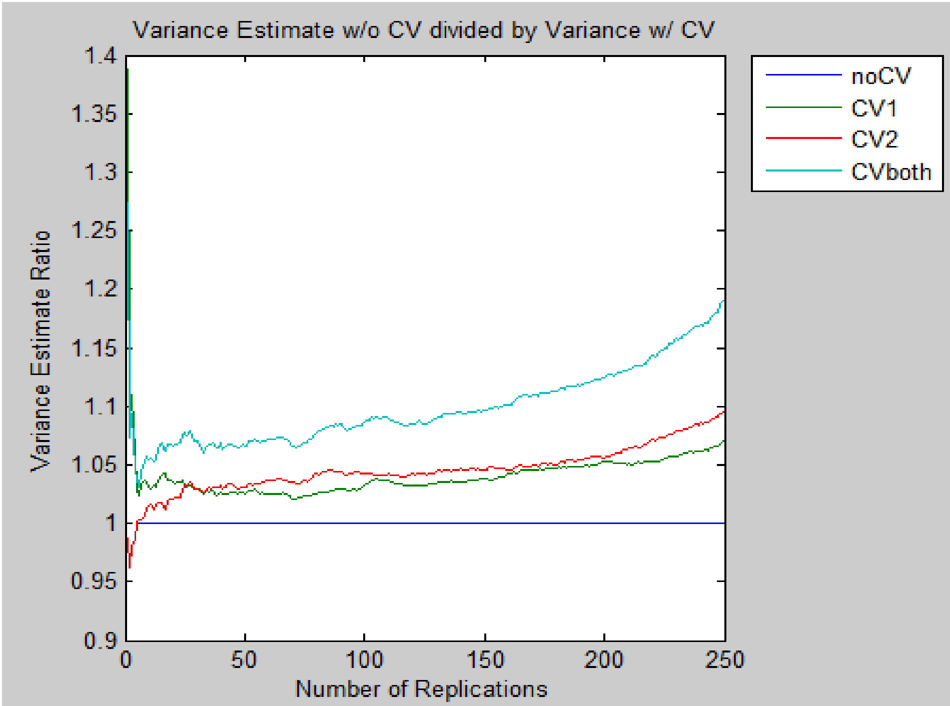


Figure 8: Variance benefit with constrained model.

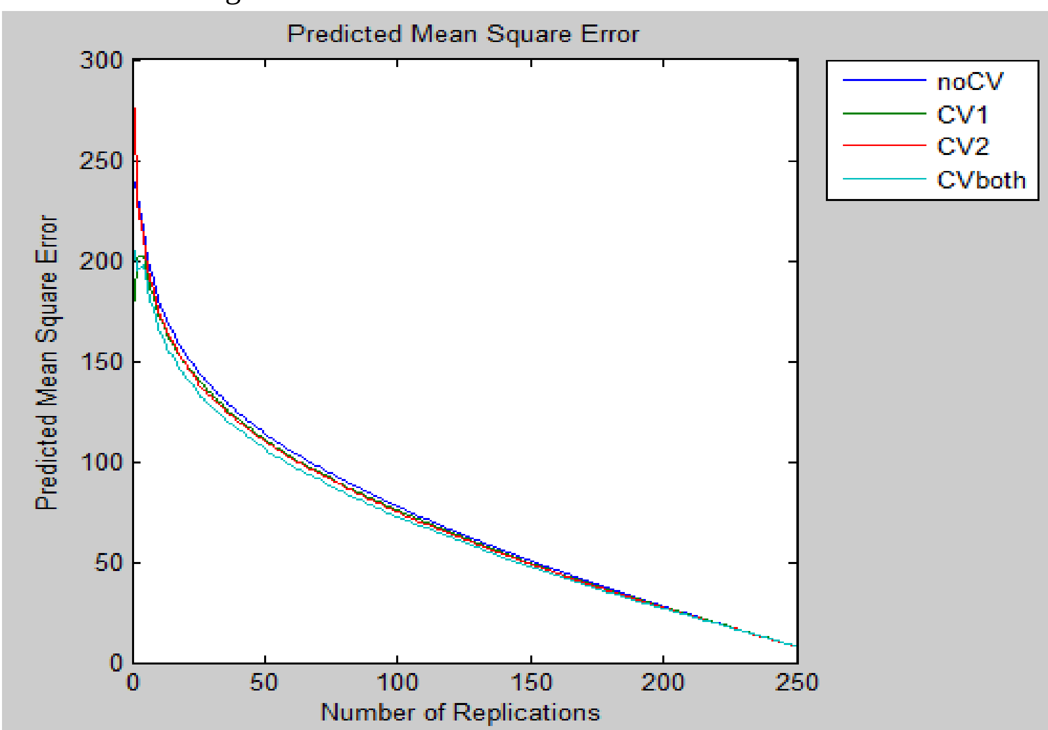


Figure 9: Prediction benefit with constrained model.

4.2.4 Summary

The traffic lane simulation can be used to make plenty of important observations about the effectiveness of control variates and optimal designs. It highlighted the important of constant covariance between the response and control variates throughout the design space as well as how the buffer size can impact the benefit of a control variate. However, most importantly, it showed that control variates interact the same each optimal criterion but not all optimal designs perform the same overall. The control variates showed the same trends through each optimal design, the appearance of one control variate to be better than another. The I-Optimal design appeared to have the lowest variance estimation and smallest half width but performed the worst on actually predicting future values. Meanwhile, the alias-optimal design showed the greatest variance estimate and largest half width but performed the best in mean square error prediction. All methods showed a slight improvement in variance estimation with control variates but were not a large enough reduction to counteract the increased degrees of freedom. The most important thing from this research is that although the initial designs did not show a benefit from control variates, the cost of checking them was so small that it did not impede the progress of creating the model. In all four designs, the time and cost of tracking the control variate was nearly zero, so when the control variate is highly correlated with the response, it will be available. And when the correlation is not high enough to have a benefit, then the analyst spent no additional resources to ensure they had the best model possible.

The modifications done to this simulation also highlight these conclusions even further. When a relationship can be established between the control variate and the response as seen in the processing time and the waiting time of a long queue, there is a clear benefit in terms of variance estimation, half-width minimization, and prediction error minimization. These are all important parts of a meta-model and any improvement is certainly a worthwhile one. Because variation is the reason for the benefit, there is indeed fluctuation from one replication to the next, which shows the importance of accounting for all possible control variates. This case showed that the processing time of the left side proved to be the beneficial control variate, while other replications showed more benefit from the right side or the combination of both. In the end, a control variate model, regardless of the optimal design used to create it, outperformed its non control variate counterpart on a regular basis.

4.3 618th TACC Simulation

4.3.1 Designs Created

The 618th TACC flight management division is concerned with the time it takes to plan a sortie and the utilization rate of their personnel. The variables, control variates, and constraints were input into JMP10, Table 25 to Table 28 show the designs that were found. Each design also includes 8 center points, making the total number of runs for the full factorial design, 40, while the optimal experimental designs are created with 30 runs.

4.3.2 Correlation

These designs were input into the ARENA Process Analyzer and the response values and control variates for each run were collected. The results were analyzed to get the correlation matrix values listed in Table 29. This matrix shows that correlation values vary considerably. However, the correlation values appear to show that the response for the time a sortie is in the planning system is much more correlated to the control variates than the utilization rate is correlated to the control variates. This difference in correlation should make the time in system response benefit much more from the inclusion of these control variates than the utilization rate. However, because of the limited time and cost, the control variates are applied to both responses.

Table 25: Design for the Full Factorial TACC model					
Full Factorial					
Run	# of FM	# Sorties	% Discrepancy Found Initially	ATM Phone Time Adjust.	In - Flight Discrepancy
1	21	210	25	0.8	35
2	23	210	25	0.8	35
3	21	240	25	0.8	35
4	23	240	25	0.8	35
5	21	210	35	0.8	35
6	23	210	35	0.8	35
7	21	240	35	0.8	35
8	23	240	35	0.8	35
9	21	210	25	1.2	35
10	23	210	25	1.2	35
11	21	240	25	1.2	35
12	23	240	25	1.2	35
13	21	210	35	1.2	35
14	23	210	35	1.2	35
15	21	240	35	1.2	35
16	23	240	35	1.2	35
17	21	210	25	0.8	40
18	23	210	25	0.8	40
19	21	240	25	0.8	40
20	23	240	25	0.8	40
21	21	210	35	0.8	40
22	23	210	35	0.8	40
23	21	240	35	0.8	40
24	23	240	35	0.8	40
25	21	210	25	1.2	40
26	23	210	25	1.2	40
27	21	240	25	1.2	40
28	23	240	25	1.2	40
29	21	210	35	1.2	40
30	23	210	35	1.2	40
31	21	240	35	1.2	40
32	23	240	35	1.2	40
33	22	225	30	1	37.5
34	22	225	30	1	37.5
35	22	225	30	1	37.5
36	22	225	30	1	37.5
37	22	225	30	1	37.5
38	22	225	30	1	37.5
39	22	225	30	1	37.5
40	22	225	30	1	37.5

Table 26: Design for the D - Optimal TACC model					
D – Optimal Design					

Design Diagnostics

D Optimal Design

D Efficiency	80.87867
G Efficiency	90.80131
A Efficiency	75.62312
Average Variance of Prediction	0.205711
Design Creation Time (seconds)	1.05

Run	# of FM	# Sorties	% Discrepancy Found Initially	ATM Phone Time Adjust.	In - Flight Discrepancy
1	21	210	25	1.2	40
2	21	210	35	0.86	35
3	21	210	35	1.2	39.5
4	21	210	25	1.2	35.5
5	21	214	25	0.8	40
6	21	236	35	1.2	35
7	21	240	25	1.14	40
8	21	240	25	0.8	35.5
9	21	240	25	1.14	35.5
10	21	240	35	0.8	35
11	21	240	35	0.8	39.5
12	23	210	25	1.2	35.5
13	23	210	25	0.86	40
14	23	210	33	1.2	40
15	23	210	35	1.2	35
16	23	215	25	0.8	35.5
17	23	214	35	0.8	39.5
18	23	235	25	1.2	40
19	23	240	25	1.14	35.5
20	23	240	25	0.8	40
21	23	240	35	0.8	35
22	23	240	35	1.14	39.5
23	22	225	30	1	37.5
24	22	225	30	1	37.5
25	22	225	30	1	37.5
26	22	225	30	1	37.5
27	22	225	30	1	37.5
28	22	225	30	1	37.5
29	22	225	30	1	37.5
30	22	225	30	1	37.5

Table 27: Design for the I - Optimal TACC model

I – Optimal Design

Design Diagnostics

I Optimal Design	
D Efficiency	79.90991
G Efficiency	88.05688
A Efficiency	76.21882
Average Variance of Prediction	0.202355
Design Creation Time (seconds)	0.916667

Run	# of FM	# Sorties	% Discrepancy Found Initially	ATM Phone Time Adjust.	In - Flight Discrepancy
1	21	210	35	0.86	35
2	21	210	35	1.2	39.5
3	21	210	25	1.2	35.5
4	21	210	25	0.854	40
5	21	214	35	0.8	39.5
6	21	215	25	0.8	35.5
7	21	237	25	1.176	40
8	21	238	35	1.17	35
9	21	240	35	0.8	39.5
10	21	240	25	0.8	35.5
11	21	240	35	1.14	39.5
12	23	210	25	0.86	35.5
13	23	210	25	1.2	40
14	23	210	35	1.2	35
15	23	210	35	1.2	39.5
16	23	214	33	0.8	40
17	23	215	35	0.8	35
18	23	236	25	1.2	35.5
19	23	240	25	0.8	40
20	23	240	25	1.14	35.5
21	23	240	35	0.8	35
22	23	240	35	1.14	39.5
23	22	225	30	1	37.5
24	22	225	30	1	37.5
25	22	225	30	1	37.5
26	22	225	30	1	37.5
27	22	225	30	1	37.5
28	22	225	30	1	37.5
29	22	225	30	1	37.5
30	22	225	30	1	37.5

Table 28: Design for the Alias Optimal TACC model

Alias Optimal Design

Design Diagnostics

I Optimal Design

D Efficiency	80.21845
G Efficiency	88.04292
A Efficiency	75.12553
Average Variance of Prediction	0.207303
Design Creation Time (seconds)	0.883333

Run	# of FM	# Sorties	% Discrepancy Found Initially	ATM Phone Time Adjust.	In - Flight Discrepancy
1	21	210	25	1.2	40
2	21	210	25	1.2	35.5
3	21	210	35	1.2	35
4	21	212	35	0.83	39.5
5	21	215	25	0.8	35.5
6	21	237	35	1.176	39.5
7	21	240	25	1.146	40
8	21	240	25	0.8	40
9	21	240	27	1.14	35
10	21	240	35	0.8	35
11	23	210	25	0.86	40
12	23	210	25	1.2	35.5
13	23	210	35	0.86	35
14	23	210	35	1.2	39.5
15	23	211	25	0.848	35.5
16	23	236	35	1.2	35
17	23	235	25.8	1.2	40
18	23	238	25	1.17	35.5
19	23	240	25	0.8	35.5
20	23	240	35	0.8	39.5
21	23	240	25	0.8	40
22	23	240	35	0.8	35
23	22	225	30	1	37.5
24	22	225	30	1	37.5
25	22	225	30	1	37.5
26	22	225	30	1	37.5
27	22	225	30	1	37.5
28	22	225	30	1	37.5
29	22	225	30	1	37.5
30	22	225	30	1	37.5

Table 29: Correlation values for all TACC designs			
Design	CV	Time in System	Utilization Rate
Full Factorial	1	-0.24153	-0.00660
	2	0.27269	0.02230
	3	0.08567	0.06367
	4	0.60294	-0.04726
	5	0.20642	0.11163
D – Optimal	1	0.23003	-0.01175
	2	0.27282	0.41532
	3	-0.04193	0.08879
	4	0.13641	-0.05981
	5	0.45406	0.19884
I – Optimal	1	0.12229	0.05721
	2	-0.00533	-0.06984
	3	0.26723	0.08341
	4	-0.02952	-0.26197
	5	-0.13283	0.22310
Alias Optimal	1	0.10733	0.12050
	2	0.13517	0.14884
	3	0.00603	0.10400
	4	0.28774	0.03380
	5	0.63550	0.41689

4.3.3 Model Creation

The next step is to develop 6 models for each design, 3 for each response. The first model is found using a stepwise function as if control variates were never tracked. This will represent the conclusions an analyst would make when ignoring control variates. The second model is the same as the first model, but significant control variates are added to the model. This approach mirrors previous research by Arnold, Nozari, and Pegden (1984) and Porta Nova and Wilson (Nov. 1989). The third model is found the same way as the first model, using the stepwise function. However, now the control variates are made available from the beginning. This represents what an analyst would conclude when tracking and applying control variates from the beginning. Additional factors may become significant or insignificant once control variates are added to the model.

4.3.4 Significant Factors

The coefficient values for the models created by JMP are shown in Table 30 for the time in system response and Table 31 for the models using utilization rate as a response. It can be easily seen by scanning the table that there are several cases for both responses where factors were added to the third model, when control variates were available for the stepwise function. However, it can also be seen that even simply adding the control variates, as seen in the second model, changes the value for each coefficient. This shows the impact control variates can have on a model. The change in coefficient means the model predicts a different value and is presenting the response surface slightly differently. The impact of additional significant factors can have a large impact on the metamodel. A factor, that was thought to have no impact on the response and was not accounted for, is now found to have an effect on the response and should be set to different values to account for this.

4.3.5 R^2adjusted Output

The R^2adjusted for all models is shown in Table 32. These values were computed by JMP when the model was found. An increase in R^2adjusted shows a model that accounts for more of the variance while also accounting for an increase in the number of terms to do so. Inspection of these values mirror the results found from inspecting the correlation values. There is a larger increase in the models using control variates when time in system is the response of interest, than there is when utilization rate is the response. In 2 of the 4 cases for utilization rate, full factorial and I-optimal, simply adding control variates to the model found without them actually decreased the R^2adjusted. Interestingly, they had the smallest correlation between control variates and utilization rate as shown in Table 29. This is a sign that the variance

Table 30: Coefficient values for each the significant figures in each TACC time in system model												
	No CV				CV Added to Initial				CV Always Available			
	Full	D	I	Alias	Full	D	I	Alias	Full	D	I	Alias
Intercept	148.57	148.51	149.49	149.30	148.19	148.09	148.87	148.38	148.21	147.56	148.89	148.25
# of FM	-0.31				-0.09						-0.19	
# Sorties										-0.62		
% Discrepancy Found	2.88	1.74	2.70	1.69	2.28	1.45	2.90	1.21	2.17	1.35	2.86	1.14
ATM Phone Time	0.03	1.23		1.92	0.17	0.89		0.83		0.93	-0.41	0.60
In - Flight Discrep.	3.03	4.30	2.40	2.94	2.80	3.67	2.83	2.65	2.72	3.79	2.87	2.68
NumFM *%DiscFnd	0.94				0.55							
NumFM * ATM	1.14				0.61						0.63	
Sortie * %DiscFnd										-0.91		
Sortie * ATM										-1.23		
%Disc * ATM												-0.84
ATM * In Flight Disc		1.14				0.96						

Table 31: Coefficient values for each the significant figures in each TACC utilization rate model												
	No CV				CV Added to Initial				CV Always Available			
	Full	D	I	Alias	Full	D	I	Alias	Full	D	I	Alias
Intercept	0.865	0.866	0.858	0.869	0.865	0.869	0.855	0.870	0.865	0.869	0.861	0.870
# of FM	-0.041	-0.041	-0.036	-0.053	-0.041	-0.037	-0.036	-0.055	-0.041	-0.036	-0.041	-0.054
# Sorties	0.055	0.051	0.050	0.069	0.055	0.051	0.049	0.072	0.055	0.051	0.050	0.070
% Discrepancy Found		0.009				0.010				0.010	0.002	0.000
ATM Phone Time											0.016	
In - Flight Discrep		0.023				0.022			0.009	0.022	0.005	
NumFM * NumSortie				-0.016				-0.014				-0.014
NumFM * %DiscFnd		0.011				0.007				0.007		
NumFM * ATM											0.017	
Sortie * %DiscFnd										0.004		-0.015
%Disc * In Flight Disc											-0.018	

explained by the control variates is not worth the extra terms needed to do so. Meanwhile, for the time in system response, adding control variates to the initial model showed improvement, and the final model also including the newly significant factors, the model's R^2adjusted increased even further. These are signs that the variance explained by the control variates is worth the increased degrees of freedom and the situation is a great candidate for the application of control variates.

Table 32: R^2adj values for each model and response

Design	Model	R^2adj for Time in System	R^2adj for Utilization Rate
Full Factorial	No CV	0.694	0.778
	CV Added to Initial	0.805	0.777
	CV Always Available	0.803	0.786
D – Optimal	No CV	0.760	0.803
	CV Added to Initial	0.849	0.835
	CV Always Available	0.875	0.842
I – Optimal	No CV	0.737	0.707
	CV Added to Initial	0.778	0.703
	CV Always Available	0.825	0.776
Alias Optimal	No CV	0.496	0.814
	CV Added to Initial	0.763	0.844
	CV Always Available	0.784	0.861

4.3.6 Minimum Variance Reduction Required

Now that the number of factors and number of control variates for each model is known, the required reduction in variance can be calculated. Table 33 shows the number of factors and control variates included in each model predicting the time in system. Table 34 shows the number of factors and control variates included in each model predicting the utilization rate. The last column of each table shows the theoretical minimum ratio between the estimated variance and the true variance in order for control variates to have a benefit due to the increased number of factors. This is found by using the ratio of $\frac{n-p-q-1}{n-p-1}$ which is the reciprocal of the values used to estimate coefficient variance. Control variates were added to the model based on their p-value calculated by the stepwise function in JMP.

Table 33: Required variance reduction for future benefits in Time in System				
Design	Model	# of Factors	# of Control Variates	$\frac{n-p-q-1}{n-p-1}$
Full Factorial	No CV	6	0	
	CV Added to Initial	6	3	0.909
	CV Always Available	2	3	0.919
D – Optimal	No CV	4	0	
	CV Added to Initial	4	5	0.800
	CV Always Available	6	5	0.783
I – Optimal	No CV	2	0	
	CV Added to Initial	2	2	0.926
	CV Always Available	5	2	0.917
Alias Optimal	No CV	3	0	
	CV Added to Initial	3	3	0.885
	CV Always Available	4	3	0.880

4.3.7 Variance Estimation and Coefficient Half Width

The models found under the various selection criteria are then input into the MATLAB code to solve for the variance estimators and the half widths of the coefficient on each significant factor, as shown in Figure 3. All potential combinations are explored but the control variate combination creating the lowest variance, including loss factor, is the model chosen for further

Table 34: Required variance reduction for future benefits in utilization rate				
Design	Model	# of Factors	# of Control Variates	Required Variance Reduction
Full Factorial	No CV	2	0	
	CV Added to Initial	2	1	0.973
	CV Always Available	3	1	0.972
D – Optimal	No CV	5	0	
	CV Added to Initial	5	1	0.958
	CV Always Available	6	1	0.957
I – Optimal	No CV	2	0	
	CV Added to Initial	2	1	0.963
	CV Always Available	7	1	0.955
Alias Optimal	No CV	3	0	
	CV Added to Initial	3	1	0.962
	CV Always Available	5	1	0.958

comparisons. The half widths for the coefficients in each of these models are shown in Table 35 for the time in system response and Table 37 for the utilization rate response. The benefit of control variates when predicting the time in system is clear. The half widths for each factor decreases from the initial model to the model with control variates added and even further improvement when other terms are allowed to enter the model after control variates are added. This is due to the decrease in the variance estimator found in Table 36. Column 4 contains the adjusted variance values used to estimate the half width. Although there is an increase in the loss factor because of the number of terms, the decrease in variance is large enough to still create a decrease in half width.

The half widths for the utilization rate models, displayed in Table 37, do not show the same trend. While the D-optimal and Alias optimal models show a minimal improvement in half width when adding control variates to the initial model, and that same improvement in the final model, the full factorial design shows no change and the I-optimal models get worse when including control variates. This is again due to the variance estimations shown in Table 38. As the number of factors included in the model increase, the estimators for half width also increase. Although there may be a decrease in the initial variance estimate shown in column 2, it does not improve more than the value determined in Table 33 which is necessary for improvement. Column 4 shows the variance once this change in the number of factors is included. There are some cases, such as the I-optimal model, where column 4 does show improvement however the half width does not. The increased number of significant factors from the final model also increases the half width. Due to solving for the simultaneous coefficient half width using the Bonferroni equation where α of each half width is $\alpha/(2p)$, increasing the number of factors also increases the half width of the coefficients. This is mostly seen in evaluating the I-Optimal design. There is a slight increase in half width even though there is a decrease in variance. This is due to the effect of the Bonferroni simultaneous confidence interval accounting for an additional 3 factors. This can make it difficult to compare the half widths of equations with a different number of significant factors, but for the purpose of this research it is an accepted effect of the approach. In large part for the utilization rate model, the reduction in variance is minimal when using control variates. As seen from the low correlation and the limited increase in R^2adjusted, this was expected. However, the change in coefficients and inclusion of additional factors may still assist in prediction ability to be seen in the next section.

Table 35: Half width values for each the significant figures in each TACC time in system model

	Full Factorial			D – Optimal			I – Optimal			Alias Optimal		
	1	2	3	1	2	3	1	2	3	1	2	3
Intercept	1.06	0.889	0.759	0.975	0.867	0.863	0.743	0.714	0.754	1.256	0.927	0.929
# of FM	1.185	0.994							0.882			
# Sorties						1.084						
% Discrepancy Found	1.185	0.994	0.849	1.176	1.046	1.03	0.883	0.848	0.895	1.507	1.112	1.114
ATM Phone Time	1.185	0.994		1.226	1.091	1.102			0.961	1.568	1.157	1.17
In - Flight Discrep.	1.185	0.994	0.849	1.249	1.112	1.1	0.968	0.931	0.984	1.599	1.18	1.184
NumFM *%DiscFnd	1.185	0.994										
NumFM * ATM	1.185	0.994							0.962			
Sortie * %DiscFnd						1.095						
Sortie * ATM						1.191						
%Disc * ATM												1.192
ATM * In Flight Disc				1.345	1.970							

1. No CV Model 2. CVs added to initial model 3. CVs always available

Table 36: Variance estimates for each time in system model			
Design	Model	$\hat{\tau}^2$	$\hat{\tau}^2 * \frac{n-p-1}{n-p-q-1}$
Full Factorial	No CV	6.695	6.695
	CV Added to Initial	4.267	4.709
	CV Always Available	4.313	4.705
D – Optimal	No CV	4.532	4.532
	CV Added to Initial	2.842	3.590
	CV Always Available	2.347	3.037
I – Optimal	No CV	3.269	3.269
	CV Added to Initial	2.761	2.992
	CV Always Available	2.572	2.817
Alias Optimal	No CV	8.103	8.103
	CV Added to Initial	3.817	4.338
	CV Always Available	3.475	3.971

Table 37: Half width values for each the significant figures in each TACC utilization model												
	Full Factorial			D – Optimal			I – Optimal			Alias Optimal		
	1	2	3	1	2	3	1	2	3	1	2	3
Intercept	0.012	0.012	0.012	0.012	0.011	0.011	0.013	0.014	0.015	0.015	0.014	0.015
# of FM	0.013	0.013	0.014	0.014	0.013	0.013	0.016	0.016	0.017	0.017	0.016	0.017
# Sorties	0.013	0.013	0.014	0.015	0.013	0.014	0.017	0.017	0.019	0.019	0.017	0.018
% Discrepancy Found				0.015	0.013	0.013			0.017			0.017
ATM Phone Time									0.019			
In - Flight Discrep.			0.014	0.016	0.014	0.015			0.019			
NumFM *%DiscFnd										0.019	0.017	0.018
NumFM * ATM				0.015	0.013	0.013						
Sortie * %DiscFnd									0.019			
Sortie * ATM						0.014						0.018
%Disc * ATM									0.020			
1. No CV Model 2. CVs added to initial model 3. CVs always available												

Table 38: Variance estimates for each utilization rate model			
Design	Model	$\hat{\tau}^2$	$\hat{\tau}^2 * \frac{n-p-1}{n-p-q-1}$
Full Factorial	No CV	0.00108	0.00108
	CV Added to Initial	0.00109	0.00112
	CV Always Available	0.00104	0.00107
D – Optimal	No CV	0.00067	0.00067
	CV Added to Initial	0.00049	0.00051
	CV Always Available	0.00050	0.00052
I – Optimal	No CV	0.00106	0.00106
	CV Added to Initial	0.00108	0.00112
	CV Always Available	0.00081	0.00085
Alias Optimal	No CV	0.00116	0.00116
	CV Added to Initial	0.00097	0.00101
	CV Always Available	0.00088	0.00092

4.3.8 Prediction

To empirically evaluate the prediction error of the different models 25 points were randomly selected throughout the design space. Each of the 3 models for each design was used to predict the “true” value of these points. The “true” value was found by running 1000 replications of the simulation at each point and accepting the average value to be true for this simulation. Table 39 shows the points that each model had to predict and the “true” response used for the prediction tests.

As suspected from the apparent correlation, large increase in R^2adjusted, and reduction in variance, the models including control variates performed better than the initial model for all designs predicting the time in system. However, despite very little correlation, minimal increase in R^2adjusted, and minimal reduction in variance, the utilization rate had mixed results. Table 40 shows the prediction improvement on the time in system response. In all designs, the prediction mean square error improved from the initial model to the model including control variates, and improved even further with the final model including all significant factors. The average prediction half width decreased for all 4 designs from the initial model to the same model including control variates, and subsequently decreased even lower when additional significant factors were added. When coverage of the true point was considered, all designs showed identical or improved coverage between the initial model and the addition of control variates. 3 of the 4 designs showed further improvement in coverage with the final model including all significant factors and control variates. The D-optimal designs decreased by only 4% or 1 of the 25 points, but still outperformed the original model by 8%.

Table 39: List of randomized points used for prediction

Test Point	Number of FMs	Number of Sorties	% Discrepancy Found	ATM Phone Time	In - Flight Discrep	Time in System	Utilization Rate
1	21	236	28.6803	0.936937	37.61127	145.49	145.49
2	21	212	34.3364	0.83507	35.06078	144.377	144.377
3	21	237	32.45836	1.13687	38.29959	147.998	147.998
4	21	228	33.98893	0.865414	37.01504	147.2	147.2
5	23	228	34.07008	1.144877	38.39736	148.212	148.212
6	23	237	29.16383	0.803442	35.56096	142.564	142.564
7	23	211	28.9718	1.051563	37.37749	144.329	144.329
8	22	215	30.02644	1.189094	35.45218	142.751	142.751
9	23	212	31.67658	1.108621	36.26107	144.457	144.457
10	21	231	32.37362	1.142462	37.62508	146.945	146.945
11	22	221	27.11859	1.075708	38.0967	144.748	144.748
12	23	220	27.25636	0.916032	35.068	141.215	141.215
13	22	224	25.72297	0.970337	35.32215	140.71	140.71
14	21	237	25.49866	1.039668	39.40799	146.289	146.289
15	23	235	30.7843	0.966333	35.10044	142.71	142.71
16	22	238	25.7521	1.097974	37.4833	143.667	143.667
17	23	221	34.36918	0.969532	39.73972	149.628	149.628
18	21	212	30.34866	1.049337	35.60402	143.151	143.151
19	23	217	33.53375	1.199405	37.62996	146.66	146.66
20	22	212	29.23848	1.197788	35.64385	142.678	142.678
21	23	216	32.60025	1.094402	35.76085	144.022	144.022
22	22	225	26.46997	1.161802	39.33959	146.272	146.272
23	21	230	27.51717	1.113466	38.76618	146.099	146.099
24	22	220	29.98879	1.198429	36.58488	144.237	144.237
25	21	224	27.11833	1.127585	35.16185	141.544	141.544

Table 40: Prediction benefits for each Time in System model				
Design	Model	Prediction MSE	Average Prediction Half Width	Coverage
Full Factorial	No CV	11.441	4.594	88%
	CV Added to Initial	8.842	3.865	88%
	CV Always Available	8.376	3.758	100%
D – Optimal	No CV	11.492	3.825	60%
	CV Added to Initial	8.534	3.441	72%
	CV Always Available	7.665	3.221	68%
I – Optimal	No CV	18.137	3.193	8%
	CV Added to Initial	12.540	3.064	24%
	CV Always Available	12.043	2.892	28%
Alias Optimal	No CV	21.230	5.084	64%
	CV Added to Initial	11.493	3.739	64%
	CV Always Available	10.318	3.599	64%

Table 41 shows the same values of prediction mean square error, average prediction half width, and coverage for the prediction of utilization rate. The prediction mean square error only decreases for the Alias-optimal design in both of the models including control variates. The full factorial model with added control variates and additional significant factors and the I-Optimal model with only adding control variates also show a decrease in prediction mean square error compared to the original model. However, for all models the coverage is very good. 100% for all models including control variates, 96% for the original Alias Optimal model, and 100% the other 3 original models. This coverage comes with a decrease of the half width in 3 of the 4 designs between the initial model and the final model, excluding the full factorial design. The full factorial and the I-Optimal models are the only models to show an increased half width when control variates are added to the original model. Both of which are increases of less than 3%. Although the estimates leading up to this step did not give it much promise on prediction ability, the addition of control variates to cover the true values was on pace with the initial model with a reduced half width even though the actual prediction error was slightly larger. In this case the initial model may already be so good that predicting the "true" value any better is very unlikely, but control variates can still offer a slight decrease in half width and still cover the "true" value.

4.3.9 Summary

Table 42 and Table 43 give a view at the percent change for all measures used in this section on the time in system response and the utilization response, respectively. As seen throughout the time in system results in Table 42, the large correlation between the control variates and the time a sortie spends being processed gave promise towards their potential benefit. This was reinforced by a large increase in R^2adjusted when the models were developed. The inclusion of control variates and additional factors meant the variance reduction must be even greater in order for it to result in a decreased half width on the coefficients. However, the variance estimators showed that this was not a problem. With correlation, an increase in explained variance, additional significant factors, and decreased coefficient half width, control variates had already offered many benefits towards this use in exploring this response surface. It proved even more important when all 4 designs showed improved prediction mean square error, reduced prediction half

Table 41: Prediction benefits for each utilization rate model

Design	Model	Prediction MSE	Average Prediction Half Width	Coverage
Full Factorial	No CV	0.00050	0.0572	100%
	CV Added to Initial	0.00052	0.0582	100%
	CV Always Available	0.00039	0.0573	100%
D – Optimal	No CV	0.00036	0.0472	100%
	CV Added to Initial	0.00042	0.0414	100%
	CV Always Available	0.00041	0.0402	100%
I – Optimal	No CV	0.00030	0.0579	100%
	CV Added to Initial	0.00022	0.0595	100%
	CV Always Available	0.00064	0.0545	100%
Alias Optimal	No CV	0.00105	0.0611	96%
	CV Added to Initial	0.00080	0.0571	100%
	CV Always Available	0.00099	0.0570	100%

width, and increased coverage when control variates were applied. This is a clear example when start to finish the benefits are clear.

The utilization rate response was not as clear. Although the correlation was small and there was only a slight increase in R^2adjusted, giving little promise for their benefit, we continued to apply the technique to investigate the results. After seeing mixed results on variance estimation and coefficient half width, we still saw a benefit in prediction. Nearly all of the designs showed a decreased half width and identical coverage. Because of the low cost and limited resources required to implement control variates, it would carry little to no cost to carry out these steps with control variates as they would be done without control variates anyways. Some of these benefits were also masked by the magnitude of the utilization rate. The percent change puts these benefits in a better light. This shows that even a small amount of correlation can lead to a small increase in variance explanation, and a marginal decrease in half width. However, because of the low cost of the technique, it can pay off in the end when the prediction half width still decreases and coverage is not compromised.

Table 42: Percent change for each Time in System model compared to the no CV model							
Design	Model	R^2adj	Intercept Coefficient HalfWidth	TauSq* Loss Factor	Prediction MSE	Coverage	Average Prediction Halfwidth
Full Factorial	CV Added to Initial	16.01%	-16.14%	-28.96%	-22.72%	0.00%	-15.88%
	CV Always Available	15.71%	-28.41%	-29.93%	-26.79%	13.64%	-18.19%
D – Optimal	CV Added to Initial	11.79%	-11.00%	-18.15%	-25.74%	20.00%	-10.05%
	CV Always Available	16.68%	-11.41%	-29.49%	-33.31%	13.33%	-15.80%
I – Optimal	CV Added to Initial	5.54%	-3.90%	-7.60%	-30.86%	200.00%	-4.05%
	CV Always Available	12.67%	1.38%	-11.44%	-33.60%	250.00%	-9.44%
Alias Optimal	CV Added to Initial	53.69%	-26.19%	-45.58%	-45.87%	0.00%	-26.45%
	CV Always Available	41.48%	-26.01%	-49.85%	-51.40%	0.00%	-29.22%

Table 43: Percent change for each Utilization Rate model compared to the no CV model							
Design	Model	R^2adj	Intercept HalfWidth	TauSq* Loss Factor	Prediction MSE	Coverage	Average Prediction Halfwidth
Full Factorial	CV Added to Initial	-0.19%	1.87%	3.79%	3.42%	0.00%	1.83%
	CV Always Available	1.15%	5.48%	-0.42%	-22.60%	0.00%	0.29%
D – Optimal	CV Added to Initial	3.91%	-12.29%	-23.16%	16.09%	0.00%	-12.43%
	CV Always Available	5.58%	-8.54%	-20.89%	11.99%	0.00%	-14.96%
I – Optimal	CV Added to Initial	-0.54%	2.88%	5.85%	-25.81%	0.00%	2.80%
	CV Always Available	10.01%	9.96%	-17.02%	116.23%	0.00%	-5.82%
Alias Optimal	CV Added to Initial	3.71%	-6.54%	-12.68%	-24.03%	4.17%	-6.57%
	CV Always Available	6.74%	-2.43%	-19.65%	-5.91%	4.17%	-6.84%

5. Conclusions and Recommendations

5.1 Chapter Overview

Control variates are an effective technique to reduce variance in metamodeling. This is true in many different scenarios, designs, and applications. However, as with any technique, the use of control variates has its own restrictions. The literature review of chapter 2 described previous research in control variates including the derivation behind their use, applications, assumptions, sources, and some alternative options. The literature review also covers metmodeling, measures of effectiveness, optimal experimental designs and analysis of covariance. All of these areas are used to show the benefits of control variates to an analyst attempting to explore the response surface of a simulation and make decisions based on the results. Control variates have been previously used by Arnold, Nozari, and Pegden (1984) on a single design, with multiple factors and multiple control variates, to gain knowledge on a single response when those control variates were added to a design. This thesis took this scenario and its known benefit and expanded it to multiple designs, multiple responses, and the addition of supplemental significant factors after control variates were added.

Chapter 3 covers the methodology used in this thesis. A combination of computer programming and analysis leads to conclusions on the application of control variates. JMP is used to create optimal experimental designs and the significant models. ARENA is used to run a simulation of the scenario we are exploring to get results of each design. MATLAB is used to analyze the results and create several of the measures of efficacy used for the results. These programs are used together to create outputs such as the R^2adjusted, correlation, mean square error variance estimation of the model, prediction mean square error, and prediction half width. These measures are used to analysis the benefit of control variates on metamodels.

Chapter 4 summarizes the results found from the traffic light and 618th TACC simulations. The traffic light simulation initially showed no significant benefits due to the parameters and factors of the simulations and designs. However, additional iterations of similar simulations showed the benefit of control variates when they are used to answer the right questions on the right simulations. The 618th TACC simulation used two responses with the same factors and control variates. This simulation highlighted the fact that the benefit of control variates is dependent on the relationship between the control variate and the response the analyst is interested in exploring.

5.2 Conclusions of Research

The conclusions of the research showed clear benefits from the use of control variates. The traffic light simulation highlighted the impact queue length can have on the effectiveness of a service time control variate compared to an interarrival control variate. The additional models took this into account showing control variates to be beneficial when the correct control variate was used on the right design space. Not all responses will benefit from the same control variates, and in this case the time in system saw significant benefits from the application of control variates, while the utilization rate saw limited benefit. However, both showed that control variates can be applied with limited cost, time, and resources and decisions can be made later on whether or not to use them without any adverse impact on the analysis.

Control variates can be applied to a wide range of scenarios. As simulations get larger and more complex there are even more candidates for control variates. However, as stated by Arnold, Nozari, and Pegden (1984) it is important to remember that although using all possibilities, the number of control variates should be limited to less than $n - p - 2$. These two examples simply show what to look for and what to expect when applying the technique for variance reduction and prediction involved with metamodeling. Control variates can be added to a simulation with very little effort. When the simulation has been completed, correlation can be checked between the control variate output and the response of interest. While this is not a binding constraint, higher correlations give higher promise of the possible future benefit. For example, the TACC simulation had control variates with correlation factors as low as .006 found to be significant when creating a metamodel. While another control variate had a correlation of .417 with the utilization rate which was not significant or beneficial. Meanwhile, the majority of significant control variates were greater than .2.

The R^2adjusted value of a model can also show the present and future impact of control variates. An increase in this value from the inclusion of control variates means the control variates explain more variance even when the decrease in error degrees of freedom is accounted for. As shown in all of the TACC time in system models, including control variates increased the R^2adjusted, and increased even further in 3 of the 4 designs when the significant factors were reexamined and the model adjusted. These trends are reflected in the prediction results. All of these models showing an increase in R^2adjusted lead to a decrease from the initial model in prediction mean square error and prediction half width. The R^2adjusted is also a great predictor of the benefit from the variance estimate. The measure of efficacy used for a coefficient half width reflects the trends of the R^2adjusted. There is a decrease from the initial model when control variates are added and further decrease when the significant factors are then adjusted. The only model to show a decrease from the second to third model also shows the smallest decrease in estimated variance.

When looking at the utilization rate of the TACC model, the 2 models that showed an increase in R^2adjusted when control variates were added showed the greatest decrease in variance estimation and the largest decrease in prediction half width. While all the models with control variates and all significant factors, model 3, showed an increase in R^2adjusted and a subsequent reduction in variance estimate and in prediction half width.

Control variates can offer several benefits to an analyst working with metamodeling. They can increase the R^2adjusted of a model, reduce the variance, reduce the coefficient half width, reduce the prediction half width, decrease the prediction mean square error, and increase the prediction coverage. However, this is when they are used properly in the correct situations. Due to their minimal cost of time and resources, any simulation with a theoretical correlation between the programmed values and the response should be checked for their impact. If the analyst sees high correlation values after the simulation has completed he should certainly include them in further steps. However, including them in a model can take almost no additional time and could be checked to be sure. If there is an increase in R^2adjusted, and additional factors become significant when the control variate is added, control variates should certainly be included in the model. This can then lead to decreased variance and half widths and a much improved metamodel. Both the model without control variates and the model with control variates could still be checked further if the analyst is not sure. Calculating the variance and coefficient half widths is an easy process to then again directly compare the two as done in this research. This comparison can easily show the benefits of control variates.

This research shows control variates can be applied to several different designs of the same simulation. While each design was constructed using different criteria and showed slightly different results, the potential benefit was still existent in each design. A design may be constructed to reduce the variance or increase the prediction ability. However, control variates can benefit the model in both of these areas even further as can be seen in the TACC designs for the time in system response.

5.3 Significance of Research

This research can be very significant to analysts working on creating a metamodel from a simulation and concerned with variance and prediction. This research supports the idea that regardless of the model used, if there is unexplained variance in the model, it could be due to other factors which can be accounted for by using control variates. By explaining this variance using control variates, a metamodel has a reduction in variance, coefficient half widths, prediction mean square error, and prediction half width. The impact of this decreased variance may come at some expense, the decrease of degrees of freedom in the model error. However, if the reduction in variance is great enough; it can lead to greater insight into the true model significant factors and their coefficients, as well as insight into the true response surface.

Control variates will not always be an effective technique for variance reduction, but this research provides the analyst with information on what to look for when analyzing the potential benefit of control variates. The low cost of setting up control variates makes it a possible technique in more situations than currently being used. If the analyst sees correlation between the control variates and response once the design has been run, he should continue the technique, and possibly continue even if limited correlation is seen. If the analyst sees a significant increase in R^2adjusted between the model with control variates and the model without them, it is further support for the analyst to use the model containing control variates. He should then check all factors to ensure no factors changed significance level. Then employing the final model should lead to a decrease in variance. This decrease will allow for smaller half widths surrounding the coefficient values and prediction estimates. All of these can be vital to an analyst when exploring the simulation and response surface.

5.4 Recommendations for Future Research

There are several potential areas for future research. This research looked at a single replication of the data in several cases. Looking at the response across several replications would allow for the estimation of the loss factor and minimum variance ratio mentioned by Porta Nova and Wilson (Nov. 1989). It would be very interesting to explore the interaction between minimum variance ratio and the prediction capability of a model with additional significant factors rather than simply the initial model with control variates as done by previously.

This research could also be expanded to more areas of application. The literature review mentions several studies which could have gained benefit from the application of control variates, as well as several studies which did not explore the additional step of creating the third model with all significant factors after control variates were added. These are areas which should be explored to show the more wide spread use of these techniques.

Antithetic variates and common random number generators are also variance reduction techniques. Further research could explore the impact of adding these techniques to those explored in this research. Comparing these techniques would show when efforts should be made to use these techniques in addition to, or instead of, control variates.

Analysis of covariance is similar to control variates. While each technique has their own assumptions and applications, further research comparing their effectiveness could be helpful to analysts contemplating the use of both techniques.

Future research could also expand on the effectiveness of using a control variate, such as interarrival time, before a queue and a control variate, such as service time, following a queue. As shown in the traffic simulation, these do change, exploring the space as to when one becomes more effective and the other less effective could also be very beneficial.

As mentioned in the thesis as replications increase, the design points can explain more variance, leaving less variance for the control variates to explain. Exploring the effectiveness of control variates as replications increase would be a very beneficial topic for future research involving control variates.

5.5 Summary

Control variates are an underappreciated technique to reduce the variance of simulation metamodels. A decrease in variance leads to decreased half widths around the coefficients of significant factors and the prediction estimates. While there are other options for variance reduction such as optimal experimental designs, antithetic variates, and common random number, control variates can be implemented for nearly no cost of time or resources as they require no additional runs and no analysis that would not be completed without them. The worst case scenario when applying control variates is that they do not have enough correlation to the response to offer any benefit and are determined insignificant, at a cost of only the time to record their values. The best case scenario can be an extreme reduction in unexplained variance. Simulations are a growing field for even more complex systems. Understanding these complex systems can be made easier when more variance is explained. The knowledge gained by the analyst and the customer can be immense, and considering it comes at the cost of simply tracking an element of the simulation, makes it very appealing to anyone analyzing a simulation regardless of the design of the experiment. This research shows control variates can make even optimal experimental designs better. A design constructed to reduce variance can have an even further reduction in variance and designs constructed for better prediction estimates can have even better prediction estimates all due to the application of control variates.

Appendix A

Matlab code for analyzing response. This will output the variance estimates and prediction mean square error and coverage.

```
% %Inputs:
% X: matrix of design,
% CV: control variates,
% Y: Response,
% T: Truth Values
% Points: Random Points X matrix

%load CCD250Face
%load FF250input
%load D250input
%load Alias250input
%load I250input


        %Solves for estimates using the average across all replications
        [n k]=size(Y); %solves for number of design points, n, and number of
                       replications
        CVBar1 = CV(:,1);
        CVBar2 = CV(:,2);
        alpha = .1;
        Transfer = [X CV Y];
        [r c] = size(X);
        DesignPoints = r/250;
        TransferNew = zeros(r,c+2+1);

for Rep = 1:250
    for ROW = 1:DesignPoints

    TransferNew((Rep-1)*(DesignPoints)+ROW,:) = Transfer(250*(ROW-1)+Rep,:);

    end
end

%Creates structure array of all potential design formats
        Xmatrices = struct('Xmatrix',[]);
       % contains data ones X1 X2 X1X1  X2X2 X1X2 CV1 CV2
Xmatrices(4).Xmatrix=TransferNew(:,1:c+2);
        %data ones X1 X2 X1X1  X2X2 X1X2 CV2
Xmatrices(3).Xmatrix=[TransferNew(:,1:c) TransferNew(:,c+2)];
        %contains data ones X1 X2 X1X1  X2X2 X1X2 CV1
Xmatrices(2).Xmatrix=TransferNew(:,1:c+1);
        %contains data ones X1 X2 X1X1  X2X2 X1X2
Xmatrices(1).Xmatrix=TransferNew(:,1:c);
      %Creates a matrix of just the design and matrix of the CV's
XmatricesNew = struct('Xmatrix',[]);
```

```
        YNew = TransferNew(:,c+3);
 %initializes variables
        BetaValue = zeros(c,4);
        HalfWidthTotal = zeros(c,4);
        BetaValueNOCV = zeros(c,250);
        BetaValueCV1 = zeros(c,250);
        BetaValueCV2 = zeros(c,250);
        BetaValueCVBoth = zeros(c,250);
        VarianceRatio2 = zeros(250,4);
        CorrelationList = zeros(250,2);
        TSQ1=zeros(1,4);
        TSQ=zeros(1,4);
        Measure1=zeros(1,4);
        Measure2=zeros(1,4);
        VarianceRatio = zeros(1,4);
        BetaValueBest = zeros(c,250);
        VarianceLF = zeros(250,4);

%loops for number of replications from 1 to 250

for Rep = 1:250

      Correlation = corr([Xmatrices(4).Xmatrix(1:DesignPoints*Rep,c+1:c+2)
YNew(1:DesignPoints*Rep,:)]);
      CorrelationList(Rep,:) = Correlation(3,1:2);


        for i=1:4
            YBar = YNew(1:DesignPoints*Rep,:);
            G=Xmatrices(i).Xmatrix(1:DesignPoints*Rep,:);
%Getting the design matrix and setting it to variable G
            betas=(G'*G)^-1*G'*YBar;      %estimate the coefficients
            gy = G*betas;                 %generate the estimates

%Implementing the estimator for the variance (tau_hat squared)
            res=YBar-gy;            %get the residuals
            tt=res'*res;            %sum of the squared residuals
        [n pq] = size(G);      %getting #coefficients p + #control variates q
            [n2 c2] = size(X);      %getting number of coefficients
            p = c2;
%renaming variables to match document where p = number jof factors
            q = pq - p;             %sets q to number of control variates
            tauSq = tt/(n-(p+q));
%MSE: sum of squares divided by DoF r-c = n-(p+q)
            VBcv = ((n-p-1)/(n-p-q-1))*tauSq*(X'*X)^-1;   %Var(BhatCV)
            M1 = ((n - p -1)/(n-(p+q)-1))*tauSq;
%Column 3 in table 5 of Nozari
            M2 = ((n - p -1)/(n-(p+q)-1))*tauSq*finv(1-alpha,6,(n-(p+q)));
%Column 4 in table 5 of Nozari
            VarBetaCV = diag(VBcv); %getting the variance of the coefficients
            HalfWidth = tinv(1-alpha/(2*(p)), n-(p)-q-1)*sqrt(VarBetaCV);
%Equation for simultaneous half width from Montgomery Intro to Linear Reg pg
98
```

```
            TSQ1(i)=tt;
            TSQ(i)=tauSq;            %Vector of MSE values
            Measure1(i)=M1;
            Measure2(i)=M2;
            BetaValue(:,i) = betas(1:p);
            HalfWidthTotal(:,i) = HalfWidth;
            VarianceRatio(i) = Measure1(1)/Measure1(i);
        end

        VarianceLF(Rep,:) = Measure1;
        OUTPUTusingAVERAGErep = [TSQ1;TSQ;Measure1;Measure2]';
        HALFWIDTHusingAVERAGErep = HalfWidthTotal;
        BetaValueNOCV(:,Rep) = BetaValue(:,1);
        BetaValueCV1(:,Rep) = BetaValue(:,2);
        BetaValueCV2(:,Rep) = BetaValue(:,3);
        BetaValueCVBoth(:,Rep) = BetaValue(:,4);
        VarianceRatio2(Rep,:) = VarianceRatio;
end

bestvariance = min(VarianceRatio2,[],2);
for i = 1:250
    if bestvariance(i) == VarianceRatio2(i,1)
        BetaValueBest(:,i) = BetaValueNOCV(:,i);
    elseif bestvariance(i) == VarianceRatio2(i,2)
        BetaValueBest(:,i) = BetaValueCV1(:,i);
    elseif bestvariance(i) == VarianceRatio2(i,3)
        BetaValueBest(:,i) = BetaValueCV2(:,i);
    else
        BetaValueBest(:,i) = BetaValueCVBoth(:,i);
    end
end

BetaValueCV = [BetaValueCV1, BetaValueCV2, BetaValueCVBoth, BetaValueBest];

B2= BetaValueNOCV;
B4=Points*B2;
[rx cx] = size(Points);
[r c]=size(B4);
B3=ones(r,c);
T2=T*ones(1,c);
B5=sqrt(B3./B4);
E=B5-T2;
MSE=(1/rx)*(diag(E'*E));

B6 = Xmatrices(1).Xmatrix*B2;
[rx cx] = size(Xmatrices(1).Xmatrix);
[r c]=size(B6);
B3=ones(r,c);
T2=YNew*ones(1,c);
E=B6-T2;
MSEmodel=(1/rx)*(diag(E'*E));

MSEnoCV=MSE;
```

```
B2=BetaValueCV;
B4=Points*B2;
[rx cx] = size(Points);
[r c]=size(B4);
B3=ones(r,c);
T2=T*ones(1,c);
B5=sqrt(B3./B4);
E=B5-T2;
MSE=(1/rx)*(diag(E'*E));
MSECV1 = MSE(1:250);  %it is only coincidence we have 2o models and 2o points
MSECV2 = MSE(251:500); %truth model data file
MSECVBoth = MSE(501:750);
MSECVBest = MSE(751:1000);

B6 = Xmatrices(1).Xmatrix*B2;
[rx cx] = size(Xmatrices(1).Xmatrix);
[r c]=size(B6);
B3=ones(r,c);
T2=YNew*ones(1,c);
E=B6-T2;
MSEmodelCV=(1/rx)*(diag(E'*E));

figure
plot(bestvariance)
hold all
plot(VarianceRatio2(:,2))
hold all
plot(VarianceRatio2(:,3))
hold all
plot(VarianceRatio2(:,4))
title('Variance Estimate w/o CV divided by Variance w/ CV');
legend('noCV','CV1','CV2','CVboth','Location','NorthEastOutside')
xlabel('Number of Replications');
ylabel('Variance Estimate Ratio');

figure
plot(abs(CorrelationList))
title('Correlation');
legend('CV1','CV2','Location','NorthEastOutside')
xlabel('Number of Replications');
ylabel('Correlation');

figure
plot(MSEnoCV)
hold all
plot(MSECV1)
hold all
plot(MSECV2)
hold all
plot(MSECVBoth)
title('Predicted Mean Square Error');
legend('noCV','CV1','CV2','CVboth','Location','NorthEastOutside')
xlabel('Number of Replications');
ylabel('Predicted Mean Square Error');
```

```
figure
plot(MSEmodel)
hold all
plot(MSEmodelCV(1:250,:))
hold all
plot(MSEmodelCV(251:500,:))
hold all
plot(MSEmodelCV(501:750,:))
hold all
plot(MSEmodelCV(751:1000,:))
title('Model Mean Square Error');
legend('noCV','CV1','CV2','CVboth','Location','NorthEastOutside')
xlabel('Number of Replications');
ylabel('Mean Square Error');

PredMSEcombined = [MSEnoCV MSECV1 MSECV2 MSECVBoth];
%ModelMSEcombined =[NoCV CV1 CV2 CVboth]
ModelMSEcombined = [MSEmodel MSEmodelCV(1:250,:) MSEmodelCV(251:500,:)
MSEmodelCV(501:750,:)];
```

Appendix B

The output from the first modification of the traffic light model for each of the 4 designs run.

Half Fraction		
Response	CV1Right	CV2Left
65.77	8.81	11.21
60.86	9.53	13.42
67.05	8.81	10.40
66.33	8.68	11.18
71.78	9.13	11.17
68.66	8.55	11.38
145.09	8.76	12.39
70.72	8.64	13.54
67.86	8.79	11.91
72.22	9.06	10.75
76.27	8.99	12.14
79.46	8.95	10.96
246.95	8.37	11.93
81.56	8.32	12.06
96.08	9.49	10.93
83.94	8.50	12.25
D-Optimal		
Response	CV1Right	CV2Left
77.835	8.843	11.37
72.802	9.925	11.495
56.796	9.387	13.123
83.237	9.157	12.272
70.658	9.049	12.732
85.731	9.007	10.708
70.201	9.295	10.993
203.57	8.842	11.104
63.078	9.533	13.573
64.983	9.336	11.423
79.905	8.773	10.419
80.455	9.171	13.143
63.759	9.254	11.409
127.355	8.787	11.185
105.121	9.055	11.331
74.288	8.704	11.251

Alias Optimal		
Response	CV1Right	CV2Left
81.295	9.535	13.032
75.586	9.077	12.253
76.492	9.033	12.845
62.872	8.858	10.745
72.296	9.295	11.035
71.721	8.806	11.139
81.104	9.443	13.522
64.295	9.417	11.554
98.096	8.83	10.546
67.756	9.099	13.191
75.27	9.27	11.369
154.432	8.681	11.189
187.702	8.833	11.247
72.807	8.687	11.251
71.045	8.819	12.009
74.52	8.542	11.542
I-Optimal		
Response	CV1Right	CV2Left
63.242	9.533	13.418
75.411	9.417	11.482
108.554	8.676	10.518
62.587	9.099	13.191
75.107	9.27	11.368
86.487	8.682	11.138
93.618	9.107	11.347
66.72	8.66	11.323
71.045	8.819	12.009
74.937	8.531	11.501
67.445	9.129	12.525
74.672	8.77	12.575
82.491	8.832	10.979
68.363	8.379	11.103
73.997	9.26	11.622
216.917	9.508	11.606

Appendix C

Quad chart

Control Variates and Optimal Designs in Metamodeling

Capt Joshua B. Meents
Advisor: Dr. Mark Friend
Reader: Dr. Ken Bauer
Department of Operational Sciences (ENS)

Research Questions

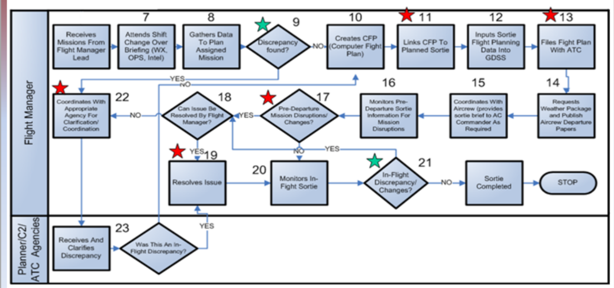
- When and how can control variates be applied?
- What are the potential benefits?
- To what designs can control variates be applied ?
- What are the hopes for improvement?

Introduction

Metamodels are often created from simulation results so the analyst can gain insight on the simulated scenario. Reducing the variance and increasing the prediction ability of these metamodels is highly sought after. Control variates (CV) can be a beneficial way to improve these metamodels. They require no additional model replications and are a very low cost approach, but they require correlation between the CV and the response of interest. Optimal designs create the most effective design based on a specific criteria while still accounting for minimum runs and specified constraints. Exploring a simulation on sortie planning shows the potential benefit of these methods when used together to create a linear regression metamodel.

Simulation Scenario

★ Used as factor of interest ★ Used as Control Variate



Summary of Control Variate Application

- Control variates show strong benefits for any design when:
  - The response of interest is highly correlated with potential control variates
  - Sufficient unexplained variance in model without control variates
  - R^2adjusted increases with the inclusion of control variates in the model
- Benefits include:
  - Reduction in model variance
  - Improvement in predicting points throughout the space
- Can be applied to any design structure
- CVs cost no additional replications or resources, when CVs are found to be not significant, the original factors can still be used

Model Results

Sortie Planning Time

Design	Model	R^2adj	Intercept Coefficient HalfWidth	TauSq* Loss Factor	Prediction MSE	Coverage	Average Prediction Halfwidth
Full Factorial	CV Added to Initial	16.01%	-16.14%	-28.96%	-22.72%	0.00%	-15.88%
	CV Always Available	15.71%	-28.41%	-29.93%	-26.79%	13.64%	-18.19%
D – Optimal	CV Added to Initial	11.79%	-11.00%	-18.15%	-25.74%	20.00%	-10.05%
	CV Always Available	16.68%	-11.41%	-29.49%	-33.31%	13.33%	-15.80%
I – Optimal	CV Added to Initial	5.54%	-3.90%	-7.60%	-30.86%	200.00%	-4.05%
	CV Always Available	12.67%	1.38%	-11.44%	-33.60%	250.00%	-9.44%
Alias Optimal	CV Added to Initial	53.69%	-26.19%	-45.58%	-45.87%	0.00%	-26.45%
	CV Always Available	41.48%	-26.01%	-49.85%	-51.40%	0.00%	-29.22%

Utilization Rate model partially met criteria to see benefits from control variate application:

- Less correlation with the CVs
- Less variance remaining to be explained

Results:

- Mixed results in all measures of effectiveness
- When R^2adj decreased, less benefit in variance and prediction
- D-Optimal and Alias Optimal Designs saw better variance and prediction measures.

Sortie Planning Time model optimal candidate for control variate application:

- High correlation with CVs
- Large amount of variance left to be explained
- Increased R^2adj for all models

Results:

- Decreased variance estimate
- Decreased coefficient half width
- Decreased mean square error of prediction
- Increased or equal coverage of prediction points
- Decreased halfwidth for prediction intervals

Flight Manager Utilization Rate

Design	Model	R^2adj	Intercept HalfWidth	TauSq* Loss Factor	Prediction MSE	Coverage	Average Prediction Halfwidth
Full Factorial	CV Added to Initial	-0.19%	1.87%	3.79%	3.42%	0.00%	1.83%
	CV Always Available	1.15%	5.48%	-0.42%	-22.60%	0.00%	0.29%
D – Optimal	CV Added to Initial	3.91%	-12.29%	-23.16%	16.09%	0.00%	-12.43%
	CV Always Available	5.58%	-8.54%	-20.89%	11.99%	0.00%	-14.96%
I – Optimal	CV Added to Initial	-0.54%	2.88%	5.85%	-25.81%	0.00%	2.80%
	CV Always Available	10.01%	9.96%	-17.02%	116.23%	0.00%	-5.82%
Alias Optimal	CV Added to Initial	3.71%	-6.54%	-12.68%	-24.03%	4.17%	-6.57%
	CV Always Available	6.74%	-2.43%	-19.65%	-5.91%	4.17%	-6.84%

Model Correlation

Design	CV	Time in System	Utilization Rate
Full Factorial	1	-0.242	-0.007
	2	0.273	0.022
	3	0.086	0.064
	4	0.603	-0.047
	5	0.206	0.112
D – Optimal	1	0.230	-0.012
	2	0.273	0.415
	3	-0.042	0.089
	4	0.136	-0.060
	5	0.454	0.199
I – Optimal	1	0.122	0.057
	2	-0.005	-0.070
	3	0.267	0.083
	4	-0.030	-0.262
	5	-0.133	0.223
Alias Optimal	1	0.107	0.121
	2	0.135	0.149
	3	0.006	0.104
	4	0.288	0.034
	5	0.636	0.417

Model Equations

- Univariate Equation: $Y_a = Y - a(C - \mu)$
- Expected Value of the Response:

$$E[Y_a] = E[Y - a(C-\mu)] = E[Y] - a\left[E[Y] - E[\mu]\right] = E[Y] - a(\mu - \mu) = E[Y]$$

- Variance of the Response:

$$\mathrm{var}(Y_a) = \mathrm{var}(Y) + a^2\,\mathrm{var}(C) - 2a\,\mathrm{cov}(Y,C)$$

- CV Multiplier:

$$a^* = \frac{\mathrm{cov}(Y,C)}{\mathrm{var}(C)}$$

γ = Factor Coefficients
$G = [X \quad (C-\mu)]$
Y = Vector of Response Values
C = Contol Variates (Vector/Matrix)
X = Factors (Vector/Matrix)
n = Number of Design Points
p = Number of Factors
q = Number of Control Variates

- Coefficients: $\hat{\gamma} = (G'G)^{-1}G'Y$
- Variance: $\hat{\tau}^2 = \dfrac{[(Y - G\hat{\gamma})'(Y - G\hat{\gamma})]}{(n-(p+q))}$

Linear Regression:

$$Y = \gamma_0 + \gamma_1 X_1 + \gamma_2 X_2 + \cdots + \gamma_p X_p + +\gamma_c'(C-\mu)$$

- Coefficient Simultaneous Confidence Interval:

$$\hat{\beta}_j \pm t_{\frac{\alpha}{2p}, n-p-q-1}\sqrt{\frac{n-p-1}{n-p-q-1}\hat{\tau}^2 (X'X)^{-1}_{jj}}$$

- Prediction Mean Square Error:

$$\left(\sum_{i=1}^{n}\left[(\hat{y}_i - y)^2\right] \Big/ n\right)$$

- Prediction Confidence Interval:

$$\hat{y} \pm t_{1-\alpha, n-p-q-1}\sqrt{\frac{n-p-1}{n-p-q-1}\hat{\tau}^2\left(1 + \left(x_0 (X'X)^{-1} x_0'\right)\right)}$$

Bibliography

Adewunmi, A., & Aickelin, U. *Investigating the Effectiveness of Variance REdcution Techniques in Manufacturing, Call Center, and Cross-Docking Discrete Event SImulation Models.* Intelligent Modelling & Analysis Research Group: School of Computer Science: The University of Nottingham.

Anonuevo, R., & Nelson, B. L. (1988). Automated Estimation and Variance Reduction Via Control Variates for Infinite-Horizon Simulations. *Computational Operations Research , Vol. 15* (No. 5), pp. 447-.

Arnold, S. F., Nozari, A., & Pegden, C. D. (1984, June). Control Variates for Multipopulation SImulation Experiments. *IIE Transactions* , 159-169.

Baesler, F. F., DaCosta, M., & Jahnsen, H. E. (2003). The Use of Simulation and Design of Experiments for Estimating Maximum Capacity in an Emergency Room. *Winter Simulation Conference.*

Bandte, O., & Mavris, D. N. (1995). Economic Uncertainty Assessment Using a Combined Design of Experiments/Monte Carlo Simulation Approach with Aplication to an HCST. *Conference of the International Society of Parametric Analysts*, (p. 9). San Diego, CA.

Bauer, K. W. (1993). Standardized Routing Variables: A New Class of Control Variates. *Journal of Statistical Computation and Simulation , 46*, 69-78.

Bauer, K., & J.R., W. (1992). Control-Variate Selection Criteria. *Naval Res. Logist. , Vol. 39*, 307-321.

Boning, D. S., & Mozumder, P. (1994). DOE/Opt: A System for Design of Experiments, Response Surface Modeling, and Optimization Using Process and Device Simulation. *Transactions on SemiConductor Manufacturing , Vol. 7* (No. 2).

Cheng, R. (1978). Analysis of SImulation Experiments under Normality Assumptions. *Journal of the Operational Research Society* , 493-97.

Chernoff, H. (1962). Optimal Design of Experiments. *Defense Technical Information Center Compilation Part Notice* .

Fort, G., & Moulines, E. *Control Variates.* LTCI, CNRS / TELECOM ParisTech.

Glynn, P. W., & Szechtman, R. (2001). *Some New Perspectives on the Method of Control Variates.* Department of Management Science and Engineering, Stanford University.

Goodman, J. (2005). Variance Reduction. Courant Institute of Mathematical Sciences, NYU.

Johnson, R. T., Lampe, T. A., & Seichter, S. (2009). Calibration of an Agent-Based Simulation Model Depicting a Refugee Camp Scenario. *Winter Simulation Conference*, (p. 8).

Jones, B., & Nachtsheim, C. (2011). Efficient Designs with Minimal Aliasing. *Technometrics , 53* (1), 62-71.

Kelton, D. W. (2000). Experimental Design for Simulation. *Winter Simulation Conference.*

Kim, S., & Henderson, S. G. (2004). *Adaptive Control Variates.* School of Operations Research and Industrial Engineering, Cornell University.

Kleijnen, J. P. (2005). An Overview of the Design and Analysis of Simulation Experiments for Sensitivity Analysis. *European Journal of Operational Research* , 287-300.

Kleijnen, J. (1974). Statistical Techniques in Simulation. *Marcel Dekker, New York* .

Kwon, C., & Tew, J. D. (1994, August). Strategies for Combining Antithetic Variates and Control Variates in Designed Simulation Experiments. *Management Science , Vol. 40* (No. 8), pp. 1021-1034.

Lavenberg, S. M. (1978). Statistical Results on COntrol Variables with Application to Queueing Network Simulation. *IBM Research Report RC7423* .

Law, A. M. (2007). *Simulation, Modeling, & Analysis* (4th ed.). Boston: McGraw-Hill.

Miller, G. A., & Chapman, J. P. (2001). Misunderstanding Analysis of Covariance. *Journal of Abnormal Psychology , 110* (1), 40-48.

Mitchell, T. J., Sacks, J., Welch, W. J., & Wynn, H. P. (1989). Design and Analysis of Computer Experiments. *Statsitical Science , Vol. 4* (No. 4), 409-435.

Montgomery, D. C., Peck, E. A., & Vining, G. G. (2006). *Introduction to Linear Regression Analysis* (4th ed.). Hoboken, NJ: John Wiley & Sons.

Montgomery, D. (2009). *Design and Analysis of Experiments* (7th ed.). Hoboken, New Jersey: John Wiley & Sons Inc.

Myers, R. H., & Montgomery, D. C. (2002). *Respnse Surface Methodology: Process and PRoduct Optimization Using Designed Experiments* (2nd ed.). New York: John Wiley & Sons Inc.

Nelson, B. L. (1990, Nov-Dec). Control Variate Remedies. *Vol. 38* (No. 6), pp. 974-992.

Nelson, B. L., & Staum, J. *Control Variates for Screening, Selection, and Estimation of the Best.* Northwestern University.

Nelson, B. L., & Yang, W.-n. (1992). *Multivariate Batch Means and Control Variates.* The Institute of Management Sciences.

Porta Nova, A. M., & Wilson, J. R. (Nov. 1989). Estimation of Multiresponse Simulation Metamodels Using Control Variates. *Management Science , Vol. 35* (No. 11), 1316-1333.

Pronzato, L., & Walter, E. (1990). Qualitative and Quantitative Experiment Design for Phenomenological Models -- A Survey. *Automatica , Vol. 26* (No. 2), 195-213.

Quinn, G., & Keough, M. (2001). Analysis of Covariance. Chapter 12.

Rubentstein, R. Y. (1981). Efficiency of Multivariate Control Variates in Monte Carlo SImulation. *Working Paper, Technion* .

Tabachnick, & Fidell. (1996). Analysis of Covariance. Chapter 8.

REPORT DOCUMENTATION PAGE		*Form Approved* *OMB No. 0704–0188*

1. REPORT DATE (DD–MM–YYYY) 21-03-2013	2. REPORT TYPE Master's Thesis	3. DATES COVERED (*From — To*) June 2012 – March 2013
4. TITLE AND SUBTITLE Control Variates and Optimal Designs in Metamodeling		5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Meents, Joshua Brian; Capt USAF		5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Air Force Institute of Technology Graduate School of Engineering and Management (AFIT) 2950 Hobson Way WPAFB OH 45433-7765		8. PERFORMING ORGANIZATION REPORT NUMBER AFIT-ENS-13-M-11
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Intentionally left blank		10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

13. SUPPLEMENTARY NOTES This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States.

14. ABSTRACT
At the heart of most modeling issues is a focus on variance reduction. Experimental designs are chosen based on both efficiency and a variety of variance based criteria. In many situations due to cost, time and availability issues it is beneficial to produce metamodels of simulations. Experimental designs for the region of operability are constructed to collect the simulation output required to construct representative metamodels. Independently, the method of control variates is a well established technique often employed to reduce variance in discrete event simulations. This thesis explores the variance reduction benefits that can be obtained by combining optimal experimental designs with control variates in multipopulation simulation experiments when constructing simulation metamodels. A variety of variance measures of effectiveness are used to demonstrate the theoretical benefits obtained by this approach. In addition, a randomly selected data set from within the design region is used to demonstrate the practical application and reduction of predictive variance obtained using this methodology.

15. SUBJECT TERMS
Control Variates, Optimal Designs, Design of Experiments, Metamodel

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON Dr. Mark Friend, ENS
a. REPORT U	b. ABSTRACT U	c. THIS PAGE U	UU	140	19b. TELEPHONE NUMBER (Include Area Code) (937)255-3636, ext 4624